

SEQUENCE LISTING

<110> BASF Plant Science GmbH
SweTree Technologies AB

5

<120> IMPROVED CONSTRUCTS FOR MARKER EXCISION BASED ON DUAL-FUNCTION
SELECTION MARKER

10 <130> PF 55443 EP

<160> 16

<170> PatentIn version 3.1

15

<210> 1

<211> 1160

<212> DNA

<213> Rhodosporidium toruloides

20

<220>

<221> CDS

<222> (1)..(1104)

<223> coding for DAAO

25

<400> 1

atg	cac	tcg	cag	aag	cgc	gtc	ggt	gtc	ctc	gga	tca	ggc	gtt	atc	ggt	48
Met	His	Ser	Gln	Lys	Arg	Val	Val	Val	Leu	Gly	Ser	Gly	Val	Ile	Gly	
1				5					10					15		

30

ctg	agc	agc	gcc	ctc	atc	ctc	gct	cgg	aag	ggc	tac	agc	gtg	cat	att	96
Leu	Ser	Ser	Ala	Leu	Ile	Leu	Ala	Arg	Lys	Gly	Tyr	Ser	Val	His	Ile	
			20					25					30			

35

ctc	gcg	cgc	gac	ttg	cgc	gag	gac	gtc	tcg	agc	cag	act	ttc	gct	tca	144
Leu	Ala	Arg	Asp	Leu	Pro	Glu	Asp	Val	Ser	Ser	Gln	Thr	Phe	Ala	Ser	
		35				40					45					

40

cca	tgg	gct	ggc	gcg	aat	tgg	acg	cct	ttc	atg	acg	ctt	aca	gac	ggt	192
Pro	Trp	Ala	Gly	Ala	Asn	Trp	Thr	Pro	Phe	Met	Thr	Leu	Thr	Asp	Gly	
	50				55				60							

45

cct	cga	caa	gca	aaa	tgg	gaa	gaa	tcg	act	ttc	aag	aag	tgg	gtc	gag	240
Pro	Arg	Gln	Ala	Lys	Trp	Glu	Glu	Ser	Thr	Phe	Lys	Lys	Trp	Val	Glu	
65				70					75					80		

	ttg gtc ccg acg ggc cat gcc atg tgg ctc aag ggg acg agg cgg ttc	288
	Leu Val Pro Thr Gly His Ala Met Trp Leu Lys Gly Thr Arg Arg Phe	
	85 90 95	
5		
	gcg cag aac gaa gac ggc ttg ctc ggg cac tgg tac aag gac atc acg	336
	Ala Gln Asn Glu Asp Gly Leu Leu Gly His Trp Tyr Lys Asp Ile Thr	
	100 105 110	
10		
	cca aat tac cgc ccc ctc cca tct tcc gaa tgt cca cct ggc gct atc	384
	Pro Asn Tyr Arg Pro Leu Pro Ser Ser Glu Cys Pro Pro Gly Ala Ile	
	115 120 125	
15		
	ggc gta acc tac gac acc ctc tcc gtc cac gca cca aag tac tgc cag	432
	Gly Val Thr Tyr Asp Thr Leu Ser Val His Ala Pro Lys Tyr Cys Gln	
	130 135 140	
20		
	tac ctt gca aga gag ctg cag aag ctc ggc gcg acg ttt gag aga cgg	480
	Tyr Leu Ala Arg Glu Leu Gln Lys Leu Gly Ala Thr Phe Glu Arg Arg	
	145 150 155 160	
25		
	acc gtt acg tcg ctt gag cag gcg ttc gac ggt gcg gat ttg gtg gtc	528
	Thr Val Thr Ser Leu Glu Gln Ala Phe Asp Gly Ala Asp Leu Val Val	
	165 170 175	
30		
	aac gct acg gga ctt ggc gcc aag tcg att gcg ggc atc gac gac caa	576
	Asn Ala Thr Gly Leu Gly Ala Lys Ser Ile Ala Gly Ile Asp Asp Gln	
	180 185 190	
35		
	gcc gcc gag cca atc cgc ggg caa acc gtc ctc gtc aag tcc cca tgc	624
	Ala Ala Glu Pro Ile Arg Gly Gln Thr Val Leu Val Lys Ser Pro Cys	
	195 200 205	
40		
	aag cga tgc acg atg gac tcg tcc gac ccc gct tct ccc gcc tac atc	672
	Lys Arg Cys Thr Met Asp Ser Ser Asp Pro Ala Ser Pro Ala Tyr Ile	
	210 215 220	
45		
	att ccc cga cca ggt ggc gaa gtc atc tgc ggc ggg acg tac ggc gtg	720
	Ile Pro Arg Pro Gly Gly Glu Val Ile Cys Gly Gly Thr Tyr Gly Val	
	225 230 235 240	
50		
	gga gac tgg gac ttg tct gtc aac cca gag acg gtc cag cgg atc ctc	768
	Gly Asp Trp Asp Leu Ser Val Asn Pro Glu Thr Val Gln Arg Ile Leu	
	245 250 255	

	aag cac tgc ttg cgc ctc gac ccg acc atc tcg agc gac gga acg atc	816
	Lys His Cys Leu Arg Leu Asp Pro Thr Ile Ser Ser Asp Gly Thr Ile	
	260 265 270	
5	gaa ggc atc gag gtc ctc cgc cac aac gtc ggc ttg cga cct gca cga	864
	Glu Gly Ile Glu Val Leu Arg His Asn Val Gly Leu Arg Pro Ala Arg	
	275 280 285	
10	cga ggc gga ccc cgc gtt gag gca gaa cgg atc gtc ctg cct ctc gac	912
	Arg Gly Gly Pro Arg Val Glu Ala Glu Arg Ile Val Leu Pro Leu Asp	
	290 295 300	
15	cgg aca aag tcg ccc ctc tcg ctc ggc agg ggc agc gca cga gcg gcg	960
	Arg Thr Lys Ser Pro Leu Ser Leu Gly Arg Gly Ser Ala Arg Ala Ala	
	305 310 315 320	
20	aag gag aag gag gtc acg ctt gtg cat gcg tat ggc ttc tcg agt gcg	1008
	Lys Glu Lys Glu Val Thr Leu Val His Ala Tyr Gly Phe Ser Ser Ala	
	325 330 335	
25	gga tac cag cag agt tgg ggc gcg gcg gag gat gtc gcg cag ctc gtc	1056
	Gly Tyr Gln Gln Ser Trp Gly Ala Ala Glu Asp Val Ala Gln Leu Val	
	340 345 350	
30	gac gag gcg ttc cag cgg tac cac ggc gcg gcg cgg gag tcg aag ttg	1104
	Asp Glu Ala Phe Gln Arg Tyr His Gly Ala Ala Arg Glu Ser Lys Leu	
	355 360 365	
35	tagggcgggg tttgtggctg tattgcgggc atctacaaga aaaaaaaaaa aaaaaa	1160
	<210> 2	
	<211> 368	
	<212> PRT	
	<213> Rhodosporidium toruloides	
40	<400> 2	
	Met His Ser Gln Lys Arg Val Val Val Leu Gly Ser Gly Val Ile Gly	
	1 5 10 15	
45	Leu Ser Ser Ala Leu Ile Leu Ala Arg Lys Gly Tyr Ser Val His Ile	
	20 25 30	
	Leu Ala Arg Asp Leu Pro Glu Asp Val Ser Ser Gln Thr Phe Ala Ser	
	35 40 45	

Pro Trp Ala Gly Ala Asn Trp Thr Pro Phe Met Thr Leu Thr Asp Gly
 50 55 60

5 Pro Arg Gln Ala Lys Trp Glu Glu Ser Thr Phe Lys Lys Trp Val Glu
 65 70 75 80

Leu Val Pro Thr Gly His Ala Met Trp Leu Lys Gly Thr Arg Arg Phe
 85 90 95

10 Ala Gln Asn Glu Asp Gly Leu Leu Gly His Trp Tyr Lys Asp Ile Thr
 100 105 110

Pro Asn Tyr Arg Pro Leu Pro Ser Ser Glu Cys Pro Pro Gly Ala Ile
 115 120 125

Gly Val Thr Tyr Asp Thr Leu Ser Val His Ala Pro Lys Tyr Cys Gln
 130 135 140

20 Tyr Leu Ala Arg Glu Leu Gln Lys Leu Gly Ala Thr Phe Glu Arg Arg
 145 150 155 160

Thr Val Thr Ser Leu Glu Gln Ala Phe Asp Gly Ala Asp Leu Val Val
 165 170 175

25 Asn Ala Thr Gly Leu Gly Ala Lys Ser Ile Ala Gly Ile Asp Asp Gln
 180 185 190

Ala Ala Glu Pro Ile Arg Gly Gln Thr Val Leu Val Lys Ser Pro Cys
 195 200 205

Lys Arg Cys Thr Met Asp Ser Ser Asp Pro Ala Ser Pro Ala Tyr Ile
 210 215 220

35 Ile Pro Arg Pro Gly Gly Glu Val Ile Cys Gly Gly Thr Tyr Gly Val
 225 230 235 240

Gly Asp Trp Asp Leu Ser Val Asn Pro Glu Thr Val Gln Arg Ile Leu
 245 250 255

40 Lys His Cys Leu Arg Leu Asp Pro Thr Ile Ser Ser Asp Gly Thr Ile
 260 265 270

Glu Gly Ile Glu Val Leu Arg His Asn Val Gly Leu Arg Pro Ala Arg
 275 280 285

45

Arg Gly Gly Pro Arg Val Glu Ala Glu Arg Ile Val Leu Pro Leu Asp
 290 295 300

5 Arg Thr Lys Ser Pro Leu Ser Leu Gly Arg Gly Ser Ala Arg Ala Ala
 305 310 315 320

Lys Glu Lys Glu Val Thr Leu Val His Ala Tyr Gly Phe Ser Ser Ala
 325 330 335

10 Gly Tyr Gln Gln Ser Trp Gly Ala Ala Glu Asp Val Ala Gln Leu Val
 340 345 350

Asp Glu Ala Phe Gln Arg Tyr His Gly Ala Ala Arg Glu Ser Lys Leu
 15 355 360 365

<210> 3
 <211> 1005
 <212> DNA
 20 <213> *Caenorhabditis elegans*
 <220>
 <221> CDS
 <222> (1)..(1002)
 25 <223> coding for DAAO
 <400> 3

atg gca aac ata att ccg aag att gca att atc ggc gaa gga gtc att 48
 Met Ala Asn Ile Ile Pro Lys Ile Ala Ile Ile Gly Glu Gly Val Ile
 30 1 5 10 15

gga tgt act tca gca ctt caa ata tca aaa gct ata cca aat gcg aaa 96
 Gly Cys Thr Ser Ala Leu Gln Ile Ser Lys Ala Ile Pro Asn Ala Lys
 20 25 30

35 ata act gtg ctc cac gat aaa cca ttt aaa aaa tcg tgc agt gca gga 144
 Ile Thr Val Leu His Asp Lys Pro Phe Lys Lys Ser Cys Ser Ala Gly
 35 40 45

40 cca gca gga tta ttt aga atc gat tat gag gag aat act gaa tac gga 192
 Pro Ala Gly Leu Phe Arg Ile Asp Tyr Glu Glu Asn Thr Glu Tyr Gly
 50 55 60

cgt gct tct ttc gcc tgg ttc tca cat ctc tat cgc act aca aaa gga 240
 45 Arg Ala Ser Phe Ala Trp Phe Ser His Leu Tyr Arg Thr Thr Lys Gly

	65		70		75		80	
5	tcc gaa acc ggc gtg aaa tta gtt tct gga cat att caa tcc gac aac Ser Glu Thr Gly Val Lys Leu Val Ser Gly His Ile Gln Ser Asp Asn	85	90	95	288			
10	ttg gag tca ttg aag caa caa caa aga gcc tat ggc gat att gtg tac Leu Glu Ser Leu Lys Gln Gln Gln Arg Ala Tyr Gly Asp Ile Val Tyr	100	105	110	336			
15	aac ttt aga ttc ttg gat gat aga gaa cgg ctg gac att ttt ccc gaa Asn Phe Arg Phe Leu Asp Asp Arg Glu Arg Leu Asp Ile Phe Pro Glu	115	120	125	384			
20	cca tca aag cac tgc att cac tac acc gcc tac gca tca gaa ggt aac Pro Ser Lys His Cys Ile His Tyr Thr Ala Tyr Ala Ser Glu Gly Asn	130	135	140	432			
25	aag tac gtg cct tat ttg aag aat ttg ctg ctt gag caa aaa atc gag Lys Tyr Val Pro Tyr Leu Lys Asn Leu Leu Leu Glu Gln Lys Ile Glu	145	150	155	480			
30	ttc aag caa caa gaa gtg acg agt ttg gac gca gtc gcc gac gct ggt Phe Lys Gln Gln Glu Val Thr Ser Leu Asp Ala Val Ala Asp Ala Gly	165	170	175	528			
35	tac gat gtt att gta aac tgc gca ggc ttg tac ggt gga aag ttg gct Tyr Asp Val Ile Val Asn Cys Ala Gly Leu Tyr Gly Gly Lys Leu Ala	180	185	190	576			
40	ggt gat gac gat act tgc tac ccc att aga gga gtc att ttg gaa gtt Gly Asp Asp Asp Thr Cys Tyr Pro Ile Arg Gly Val Ile Leu Glu Val	195	200	205	624			
45	gat gca cca tgg cac aag cac ttc aat tat cga gac ttt act act ttc Asp Ala Pro Trp His Lys His Phe Asn Tyr Arg Asp Phe Thr Thr Phe	210	215	220	672			
50	aca att cca aaa gag cac agc gtg gtg gtt ggg tcc acc aag cag gac Thr Ile Pro Lys Glu His Ser Val Val Val Gly Ser Thr Lys Gln Asp	225	230	235	720			
55	aat cga tgg gat ttg gag atc acc gac gag gat aga aat gat att ttg Asn Arg Trp Asp Leu Glu Ile Thr Asp Glu Asp Arg Asn Asp Ile Leu	245	250	255	768			

aaa cga tac att gct tta cat cct gga atg aga gag cca aag att atc 816
 Lys Arg Tyr Ile Ala Leu His Pro Gly Met Arg Glu Pro Lys Ile Ile
 260 265 270

5
 aaa gaa tgg tca gca ctt cgc ccg gga cgt aag cat gtc aga att gaa 864
 Lys Glu Trp Ser Ala Leu Arg Pro Gly Arg Lys His Val Arg Ile Glu
 275 280 285

10
 gcg cag aag agg aca tct gtt gga aac tca aaa gat tat atg gtt gtg 912
 Ala Gln Lys Arg Thr Ser Val Gly Asn Ser Lys Asp Tyr Met Val Val
 290 295 300

15
 cat cac tat ggt cac ggg agc aac gga ttc acg ttg ggt tgg gga aca 960
 His His Tyr Gly His Gly Ser Asn Gly Phe Thr Leu Gly Trp Gly Thr
 305 310 315 320

20
 gca att gaa gca act aaa ctt gtt aag act gca cta gga tta taa 1005
 Ala Ile Glu Ala Thr Lys Leu Val Lys Thr Ala Leu Gly Leu
 325 330

<210> 4
 <211> 334
 <212> PRT
 25 <213> *Caenorhabditis elegans*

<400> 4
 Met Ala Asn Ile Ile Pro Lys Ile Ala Ile Ile Gly Glu Gly Val Ile .
 1 5 10 15

30
 Gly Cys Thr Ser Ala Leu Gln Ile Ser Lys Ala Ile Pro Asn Ala Lys
 20 25 30

35
 Ile Thr Val Leu His Asp Lys Pro Phe Lys Lys Ser Cys Ser Ala Gly
 35 40 45

Pro Ala Gly Leu Phe Arg Ile Asp Tyr Glu Glu Asn Thr Glu Tyr Gly
 50 55 60

40
 Arg Ala Ser Phe Ala Trp Phe Ser His Leu Tyr Arg Thr Thr Lys Gly
 65 70 75 80

Ser Glu Thr Gly Val Lys Leu Val Ser Gly His Ile Gln Ser Asp Asn
 85 90 95

45

10

15

20

25

30

35

40

45

Ala Ile Glu Ala Thr Lys Leu Val Lys Thr Ala Leu Gly Leu
 325 330'

<210> 5
 <211> 1186
 <212> DNA
 <213> Nectria haematococca

5

<220>
 <221> CDS
 <222> (42)..(1124)
 <223> coding for DAAO

10

<400> 5
 agcgacttga atttagcgaa aagaacttgt caaccacaat c atg tcc aac aca atc 56
 Met Ser Asn Thr Ile
 1 5

15

gtc gtc gtt ggt gcc ggt gtc att ggc ttg acg tcg gcc ttg ttg ctc 104
 Val Val Val Gly Ala Gly Val Ile Gly Leu Thr Ser Ala Leu Leu Leu
 10 15 20

20

tcc aag aac aag ggc aac aag atc acc gtc gtg gcc aag cac atg ccc 152
 Ser Lys Asn Lys Gly Asn Lys Ile Thr Val Val Ala Lys His Met Pro
 25 30 35

25

ggc gac tat gac gtt gaa tac gcc tcg cct ttt gct ggt gcc aac cac 200
 Gly Asp Tyr Asp Val Glu Tyr Ala Ser Pro Phe Ala Gly Ala Asn His
 40 45 50

30

tcc ccc atg gcg acg gaa gag agc agc gaa tgg gaa cgt cgc act tgg 248
 Ser Pro Met Ala Thr Glu Glu Ser Ser Glu Trp Glu Arg Arg Thr Trp
 55 60 65

35

tac gag ttt aag aga ctg gtc gag gag gtc cct gag gcc ggt gtt cat 296
 Tyr Glu Phe Lys Arg Leu Val Glu Glu Val Pro Glu Ala Gly Val His
 70 75 80 85

ttc cag aag tct cgc atc cag agg cgc aat gtg gac act gaa aag gcg 344
 Phe Gln Lys Ser Arg Ile Gln Arg Arg Asn Val Asp Thr Glu Lys Ala
 90 95 100

40

cag agg tct ggt ttc cca gac gcc ctc ttc tcg aaa gaa ccc tgg ttc 392
 Gln Arg Ser Gly Phe Pro Asp Ala Leu Phe Ser Lys Glu Pro Trp Phe
 105 110 115

45

aag aac atg ttt gag gac ttc cgt gag cag cac cct agc gag gtc atc 440
 Lys Asn Met Phe Glu Asp Phe Arg Glu Gln His Pro Ser Glu Val Ile

	120	125	130	
	ccc ggt tac gac tct ggc tgc gag ttc aca tcg gtg tgc atc aac acg			488
	Pro Gly Tyr Asp Ser Gly Cys Glu Phe Thr Ser Val Cys Ile Asn Thr			
5	135	140	145	
	gcc atc tac ctc ccc tgg ctc ctc ggc cag tgc atc aag aat ggc gtc			536
	Ala Ile Tyr Leu Pro Trp Leu Leu Gly Gln Cys Ile Lys Asn Gly Val			
10	150	155	160	165
	atc gtc aag cgc gcc atc ctc aac gac att agc gag gcc aag aag ctg			584
	Ile Val Lys Arg Ala Ile Leu Asn Asp Ile Ser Glu Ala Lys Lys Leu			
	170	175	180	
15	agc cac gcg ggc aag acg ccc aat atc atc gtc aac gcc acg ggt ctc			632
	Ser His Ala Gly Lys Thr Pro Asn Ile Ile Val Asn Ala Thr Gly Leu			
	185	190	195	
	ggc tcc tac aag ctg ggc ggt gtc gag gac aag acc atg gcg cct gcg			680
20	Gly Ser Tyr Lys Leu Gly Gly Val Glu Asp Lys Thr Met Ala Pro Ala			
	200	205	210	
	cgg gga cag att gtg gtt gtg cgc aac gag agc agc ccc atg ctc ctc			728
	Arg Gly Gln Ile Val Val Val Arg Asn Glu Ser Ser Pro Met Leu Leu			
25	215	220	225	
	act tca ggt gtc gag gac ggc ggt gct gat gtc atg tac ttg atg cag			776
	Thr Ser Gly Val Glu Asp Gly Gly Ala Asp Val Met Tyr Leu Met Gln			
30	230	235	240	245
	cga gca gct ggc ggt ggc acc atc ctg ggc ggt acc tac gac gtt ggc			824
	Arg Ala Ala Gly Gly Gly Thr Ile Leu Gly Gly Thr Tyr Asp Val Gly			
	250	255	260	
35	aac tgg gag tct cag cca gac ccc aac atc gcg aat cgc atc atg cag			872
	Asn Trp Glu Ser Gln Pro Asp Pro Asn Ile Ala Asn Arg Ile Met Gln			
	265	270	275	
	cgc atc gtc gag gtg cgg ccc gag att gcc aac ggc aag ggc gtc aag			920
40	Arg Ile Val Glu Val Arg Pro Glu Ile Ala Asn Gly Lys Gly Val Lys			
	280	285	290	
	ggg ctg agc gtg atc cga cac gcc gtc ggc atg cgg ccg tgg cga aag			968
	Gly Leu Ser Val Ile Arg His Ala Val Gly Met Arg Pro Trp Arg Lys			
45	295	300	305	

gac gga gtc agg atc gag gag gag aag ctg gat gat gag act tgg atc 1016
 Asp Gly Val Arg Ile Glu Glu Glu Lys Leu Asp Asp Glu Thr Trp Ile
 310 315 320 325
 5
 gtg cac aac tac gga cac tct gga tgg ggt tac cag ggt tcg tat ggt 1064
 Val His Asn Tyr Gly His Ser Gly Trp Gly Tyr Gln Gly Ser Tyr Gly
 330 335 340
 10
 tgt gct gag aat gta gtc cag ttg gtt gac aag gtc ggc aag gcg gcc 1112
 Cys Ala Glu Asn Val Val Gln Leu Val Asp Lys Val Gly Lys Ala Ala
 345 350 355
 aag tct aag ctg tagttgaaaa ggcctgaatg agtaatagta attggatatt 1164
 15
 Lys Ser Lys Leu
 360
 ggaaataaccg tatttgccct cg 1186
 20
 <210> 6
 <211> 361
 <212> PRT
 <213> Nectria haematococca
 25
 <400> 6
 Met Ser Asn Thr Ile Val Val Val Gly Ala Gly Val Ile Gly Leu Thr
 1 5 10 15
 Ser Ala Leu Leu Leu Ser Lys Asn Lys Gly Asn Lys Ile Thr Val Val
 30
 20 25 30
 Ala Lys His Met Pro Gly Asp Tyr Asp Val Glu Tyr Ala Ser Pro Phe
 35 40 45
 35
 Ala Gly Ala Asn His Ser Pro Met Ala Thr Glu Glu Ser Ser Glu Trp
 50 55 60
 Glu Arg Arg Thr Trp Tyr Glu Phe Lys Arg Leu Val Glu Glu Val Pro
 65 70 75 80
 40
 Glu Ala Gly Val His Phe Gln Lys Ser Arg Ile Gln Arg Arg Asn Val
 85 90 95
 Asp Thr Glu Lys Ala Gln Arg Ser Gly Phe Pro Asp Ala Leu Phe Ser
 45
 100 105 110

10

15

20

25

30

35

40

45

Asp Glu Thr Trp Ile Val His Asn Tyr Gly His Ser Gly Trp Gly Tyr
325 330 335

Gln Gly Ser Tyr Gly Cys Ala Glu Asn Val Val Gln Leu Val Asp Lys
 340 345 350

5 Val Gly Lys Ala Ala Lys Ser Lys Leu
 355 360

<210> 7

<211> 1071

<212> DNA

10 <213> *Trigonopsis variabilis*

<220>

<221> CDS

<222> (1)..(1068)

15 <223>

<400> 7

20 atg gct aaa atc gtt gtt att ggt gcc ggt gtt gcc ggt tta act aca 48
 Met Ala Lys Ile Val Val Ile Gly Ala Gly Val Ala Gly Leu Thr Thr
 1 5 10 15

25 gct ctt caa ctt ctt cgt aaa gga cat gag gtt aca att gtg tcc gag 96
 Ala Leu Gln Leu Leu Arg Lys Gly His Glu Val Thr Ile Val Ser Glu
 20 25 30

30 ttt acg ccc ggt gat ctt agt atc gga tat acc tcg cct tgg gca ggt 144
 Phe Thr Pro Gly Asp Leu Ser Ile Gly Tyr Thr Ser Pro Trp Ala Gly
 35 40 45

30 gcc aac tgg ctc aca ttt tac gat gga ggc aag tta gcc gac tac gat 192
 Ala Asn Trp Leu Thr Phe Tyr Asp Gly Gly Lys Leu Ala Asp Tyr Asp
 50 55 60

35 gcc gtc tct tat cct atc ttg cga gag ctg gct cga agc agc ccc gag 240
 Ala Val Ser Tyr Pro Ile Leu Arg Glu Leu Ala Arg Ser Ser Pro Glu
 65 70 75 80

40 gct gga att cga ctc atc agc caa cgc tcc cat gtt ctc aag cgt gat 288
 Ala Gly Ile Arg Leu Ile Ser Gln Arg Ser His Val Leu Lys Arg Asp
 85 90 95

45 ctt cct aaa ctg gaa gtt gcc atg tcg gcc atc tgt caa cgc aat ccc 336
 Leu Pro Lys Leu Glu Val Ala Met Ser Ala Ile Cys Gln Arg Asn Pro
 100 105 110

	tgg	ttc	aaa	aac	aca	gtc	gat	tct	ttc	gag	att	atc	gag	gac	agg	tcc	384
	Trp	Phe	Lys	Asn	Thr	Val	Asp	Ser	Phe	Glu	Ile	Ile	Glu	Asp	Arg	Ser	
			115					120							125		
5	agg	att	gtc	cac	gat	gat	gtg	gct	tat	cta	gtc	gaa	ttt	cgt	tcc	gtt	432
	Arg	Ile	Val	His	Asp	Asp	Val	Ala	Tyr	Leu	Val	Glu	Phe	Arg	Ser	Val	
			130					135					140				
10	tgt	atc	cac	acc	gga	gtc	tac	ttg	aac	tgg	ctg	atg	tcc	caa	tgc	tta	480
	Cys	Ile	His	Thr	Gly	Val	Tyr	Leu	Asn	Trp	Leu	Met	Ser	Gln	Cys	Leu	
	145					150					155					160	
15	tcg	ctc	ggc	gcc	acg	gtg	gtt	aaa	cgt	cga	gtg	aac	cat	atc	aag	gat	528
	Ser	Leu	Gly	Ala	Thr	Val	Val	Lys	Arg	Arg	Val	Asn	His	Ile	Lys	Asp	
					165					170					175		
20	gcc	aat	tta	cta	cac	tcc	tca	gga	tca	cgc	ccc	gac	gtg	att	gtc	aac	576
	Ala	Asn	Leu	Leu	His	Ser	Ser	Gly	Ser	Arg	Pro	Asp	Val	Ile	Val	Asn	
					180					185					190		
25	tgt	agt	ggt	ctc	ttt	gcc	cgg	ttc	ttg	gga	ggc	gtc	gag	gac	aag	aag	624
	Cys	Ser	Gly	Leu	Phe	Ala	Arg	Phe	Leu	Gly	Gly	Val	Glu	Asp	Lys	Lys	
			195					200						205			
30	atg	tac	cct	att	cga	gga	caa	gtc	gtc	ctt	gtt	cga	aac	tct	ctt	cct	672
	Met	Tyr	Pro	Ile	Arg	Gly	Gln	Val	Val	Leu	Val	Arg	Asn	Ser	Leu	Pro	
			210					215					220				
35	ttt	atg	gcc	tcc	ttt	tcc	agc	act	cct	gaa	aaa	gaa	aat	gaa	gac	gaa	720
	Phe	Met	Ala	Ser	Phe	Ser	Ser	Thr	Pro	Glu	Lys	Glu	Asn	Glu	Asp	Glu	
	225					230					235					240	
40	gct	cta	tat	atc	atg	acc	cga	ttc	gat	ggc	act	tct	atc	att	ggc	ggc	768
	Ala	Leu	Tyr	Ile	Met	Thr	Arg	Phe	Asp	Gly	Thr	Ser	Ile	Ile	Gly	Gly	
					245					250					255		
45	tgt	ttc	caa	ccc	aac	aac	tgg	tca	tcc	gaa	ccc	gat	cct	tct	ctc	acc	816
	Cys	Phe	Gln	Pro	Asn	Asn	Trp	Ser	Ser	Glu	Pro	Asp	Pro	Ser	Leu	Thr	
				260						265					270		
50	cat	cga	atc	ctg	tct	aga	gcc	ctc	gac	cga	ttc	ccg	gaa	ctg	acc	aaa	864
	His	Arg	Ile	Leu	Ser	Arg	Ala	Leu	Asp	Arg	Phe	Pro	Glu	Leu	Thr	Lys	
				275				280						285			

gat ggc cct ctt gac att gtg cgc gaa tgc gtt ggc cac cgt cct ggt 912
 Asp Gly Pro Leu Asp Ile Val Arg Glu Cys Val Gly His Arg Pro Gly
 290 295 300

5 aga gag ggc ggt ccc cga gta gaa tta gag aag atc ccc ggc gtt ggc 960
 Arg Glu Gly Gly Pro Arg Val Glu Leu Glu Lys Ile Pro Gly Val Gly
 305 310 315 320

10 ttt gtt gtc cat aac tat ggt gcc gcc ggt gct ggt tac caa tcc tct 1008
 Phe Val Val His Asn Tyr Gly Ala Ala Gly Ala Gly Tyr Gln Ser Ser
 325 330 335

15 tac ggc atg gct gat gaa gct gtt tct tac gtc gaa aga gct ctt act 1056
 Tyr Gly Met Ala Asp Glu Ala Val Ser Tyr Val Glu Arg Ala Leu Thr
 340 345 350

cgt cca aac ctt tag 1071
 Arg Pro Asn Leu
 355

20

<210> 8
 <211> 356
 <212> PRT
 25 <213> Trigonopsis variabilis

<400> 8
 Met Ala Lys Ile Val Val Ile Gly Ala Gly Val Ala Gly Leu Thr Thr
 1 5 10 15

30 Ala Leu Gln Leu Leu Arg Lys Gly His Glu Val Thr Ile Val Ser Glu
 20 25 30

35 Phe Thr Pro Gly Asp Leu Ser Ile Gly Tyr Thr Ser Pro Trp Ala Gly
 35 40 45

Ala Asn Trp Leu Thr Phe Tyr Asp Gly Gly Lys Leu Ala Asp Tyr Asp
 50 55 60

40 Ala Val Ser Tyr Pro Ile Leu Arg Glu Leu Ala Arg Ser Ser Pro Glu
 65 70 75 80

Ala Gly Ile Arg Leu Ile Ser Gln Arg Ser His Val Leu Lys Arg Asp
 85 90 95

45

Leu Pro Lys Leu Glu Val Ala Met Ser Ala Ile Cys Gln Arg Asn Pro
100 105 110

5 Trp Phe Lys Asn Thr Val Asp Ser Phe Glu Ile Ile Glu Asp Arg Ser
115 120 125

Arg Ile Val His Asp Asp Val Ala Tyr Leu Val Glu Phe Arg Ser Val
130 135 140

10 Cys Ile His Thr Gly Val Tyr Leu Asn Trp Leu Met Ser Gln Cys Leu
145 150 155 160

Ser Leu Gly Ala Thr Val Val Lys Arg Arg Val Asn His Ile Lys Asp
165 170 175

15 Ala Asn Leu Leu His Ser Ser Gly Ser Arg Pro Asp Val Ile Val Asn
180 185 190

20 Cys Ser Gly Leu Phe Ala Arg Phe Leu Gly Gly Val Glu Asp Lys Lys
195 200 205

Met Tyr Pro Ile Arg Gly Gln Val Val Leu Val Arg Asn Ser Leu Pro
210 215 220

25 Phe Met Ala Ser Phe Ser Ser Thr Pro Glu Lys Glu Asn Glu Asp Glu
225 230 235 240

Ala Leu Tyr Ile Met Thr Arg Phe Asp Gly Thr Ser Ile Ile Gly Gly
245 . 250 255

30 Cys Phe Gln Pro Asn Asn Trp Ser Ser Glu Pro Asp Pro Ser Leu Thr
260 265 270

His Arg Ile Leu Ser Arg Ala Leu Asp Arg Phe Pro Glu Leu Thr Lys
35 275 280 285

Asp Gly Pro Leu Asp Ile Val Arg Glu Cys Val Gly His Arg Pro Gly
290 295 300

40 Arg Glu Gly Gly Pro Arg Val Glu Leu Glu Lys Ile Pro Gly Val Gly
305 310 315 320

Phe Val Val His Asn Tyr Gly Ala Ala Gly Ala Gly Tyr Gln Ser Ser
325 330 335

45

Tyr Gly Met Ala Asp Glu Ala Val Ser Tyr Val Glu Arg Ala Leu Thr
 340 345 350

5 Arg Pro Asn Leu
 355

<210> 9

<211> 1047

10 <212> DNA

<213> Schizosaccharomyces pombe

<220>

<221> CDS

15 <222> (22)..(1041)

<223> coding for DAAO

<400> 9

20 atgactaagg aaaataagcc a aga gat att gtc atc gtt ggc gct ggc gtt 51
 Arg Asp Ile Val Ile Val Gly Ala Gly Val
 1 5 10

25 att gga ttg acc act gct tgg att ctt tca gac ttg ggt ctt gct cct 99
 Ile Gly Leu Thr Thr Ala Trp Ile Leu Ser Asp Leu Gly Leu Ala Pro
 15 20 25

30 cgt att aag gtg att gcc aag tat acg cct gaa gat cgt tct gta gaa 147
 Arg Ile Lys Val Ile Ala Lys Tyr Thr Pro Glu Asp Arg Ser Val Glu
 30 35 40

tac act tcc cct tgg gct ggc gca aat ttc tgt agc att tct gct act 195
 Tyr Thr Ser Pro Trp Ala Gly Ala Asn Phe Cys Ser Ile Ser Ala Thr
 45 50 55

35 gat gac aat gct ttg cgc tgg gat aaa atc act tac cat cgt ttc gcc 243
 Asp Asp Asn Ala Leu Arg Trp Asp Lys Ile Thr Tyr His Arg Phe Ala
 60 65 70

40 tac ttg gcg aaa act cgt cct gaa gca gga atc cgt ttt gct gat ctt 291
 Tyr Leu Ala Lys Thr Arg Pro Glu Ala Gly Ile Arg Phe Ala Asp Leu
 75 80 85 90

45 cga gaa ttg tgg gag tac gag ccg aaa cac gac aaa atc aga tcc tgg 339
 Arg Glu Leu Trp Glu Tyr Glu Pro Lys His Asp Lys Ile Arg Ser Trp
 95 100 105

	aat acc tat gtc aga gat ttc aaa gtt atc cct gaa aaa gat ctt cca	387
	Asn Thr Tyr Val Arg Asp Phe Lys Val Ile Pro Glu Lys Asp Leu Pro	
	110 115 120	
5	gga gaa tgt atc tac gga cat aag gcc acc acc ttt tta atc aac gct	435
	Gly Glu Cys Ile Tyr Gly His Lys Ala Thr Thr Phe Leu Ile Asn Ala	
	125 130 135	
10	cct cat tac ttg aat tat atg tac aag ctg ctc att gaa gct ggc gtc	483
	Pro His Tyr Leu Asn Tyr Met Tyr Lys Leu Leu Ile Glu Ala Gly Val	
	140 145 150	
15	gaa ttt gaa aag aaa gaa ttg agt cac atc aaa gag act gtc gaa gaa	531
	Glu Phe Glu Lys Lys Glu Leu Ser His Ile Lys Glu Thr Val Glu Glu	
	155 160 165 170	
20	act cca gaa gct tca gta gta ttt aat tgc act ggt ctc tgg gct tcc	579
	Thr Pro Glu Ala Ser Val Val Phe Asn Cys Thr Gly Leu Trp Ala Ser	
	175 180 185	
25	aaa ttg ggt ggc gtt gaa gac ccg gac gtt tat ccg act cgt gga cat	627
	Lys Leu Gly Gly Val Glu Asp Pro Asp Val Tyr Pro Thr Arg Gly His	
	190 195 200	
30	ggt gtt ttg gtt aag gct cct cat gta aca gaa act cgc att ttg aat	675
	Val Val Leu Val Lys Ala Pro His Val Thr Glu Thr Arg Ile Leu Asn	
	205 210 215	
35	ggc aag aac tct gat acc tat att att cct cgt ccc tta aat ggt gga	723
	Gly Lys Asn Ser Asp Thr Tyr Ile Ile Pro Arg Pro Leu Asn Gly Gly	
	220 225 230	
40	gtc att tgc ggc ggt ttc atg caa cca gga aac tgg gat cgt gaa att	771
	Val Ile Cys Gly Gly Phe Met Gln Pro Gly Asn Trp Asp Arg Glu Ile	
	235 240 245 250	
45	cac cct gaa gac act ttg gat atc ctt aag aga aca tcg gct ttg atg	819
	His Pro Glu Asp Thr Leu Asp Ile Leu Lys Arg Thr Ser Ala Leu Met	
	255 260 265	
50	cca gaa ttg ttc cac ggc aag ggt ccg gag ggt gct gaa att att caa	867
	Pro Glu Leu Phe His Gly Lys Gly Pro Glu Gly Ala Glu Ile Ile Gln	
	270 275 280	

gaa tgt gtc gga ttc cgt cct tct cga aag ggt ggt gcc cgc gta gag 915
 Glu Cys Val Gly Phe Arg Pro Ser Arg Lys Gly Gly Ala Arg Val Glu
 285 290 295

5 ctt gat gtt gtt ccc ggc acc tca gtc ccc ctt gtt cat gat tac ggt 963
 Leu Asp Val Val Pro Gly Thr Ser Val Pro Leu Val His Asp Tyr Gly
 300 305 310

10 gct tct ggc aca gga tac caa gct ggt tat ggt atg gct ctt gac tct 1011
 Ala Ser Gly Thr Gly Tyr Gln Ala Gly Tyr Gly Met Ala Leu Asp Ser
 315 320 325 330

gtc atg ttg gct ctt cct aaa atc aaa ttg gcttag 1047
 Val Met Leu Ala Leu Pro Lys Ile Lys Leu
 15 335 340

<210> 10
 <211> 340
 20 <212> PRT
 <213> Schizosaccharomyces pombe

<400> 10
 Arg Asp Ile Val Ile Val Gly Ala Gly Val Ile Gly Leu Thr Thr Ala
 25 1 5 10 15
 Trp Ile Leu Ser Asp Leu Gly Leu Ala Pro Arg Ile Lys Val Ile Ala
 20 25 30
 30 Lys Tyr Thr Pro Glu Asp Arg Ser Val Glu Tyr Thr Ser Pro Trp Ala
 35 40 45
 Gly Ala Asn Phe Cys Ser Ile Ser Ala Thr Asp Asp Asn Ala Leu Arg
 50 55 60
 35 Trp Asp Lys Ile Thr Tyr His Arg Phe Ala Tyr Leu Ala Lys Thr Arg
 65 70 75 80
 Pro Glu Ala Gly Ile Arg Phe Ala Asp Leu Arg Glu Leu Trp Glu Tyr
 40 85 90 95
 Glu Pro Lys His Asp Lys Ile Arg Ser Trp Asn Thr Tyr Val Arg Asp
 100 105 110
 45

<210> 11
 <211> 963
 <212> DNA
 <213> Streptomyces coelicolor

5

<220>
 <221> CDS
 <222> (31)..(957)
 <223> coding for DAAO

10

<220>
 <221> misc_feature
 <222> (880)..(936)
 <223> DAAO signature

15

<400> 11
 gtggaaaccg aactggatga cgagcgggat ggc gaa gtc gtc gtg gtc ggc ggc 54
 Gly Glu Val Val Val Val Gly Gly
 1 5

20

ggg gtg atc ggg ctg acg acg gcc gtc gtc ctc gcc gag cgg ggc aga 102
 Gly Val Ile Gly Leu Thr Thr Ala Val Val Leu Ala Glu Arg Gly Arg
 10 15 20

25

cgg gtg cgg ctg tgg acc cgg gag ccc gcg gag cgg acc acc tcg gtg 150
 Arg Val Arg Leu Trp Thr Arg Glu Pro Ala Glu Arg Thr Thr Ser Val
 25 30 35 40

30

gta gcg ggc ggg ctg tgg tgg ccg tac cgg atc gag ccg gtc gcg ctg 198
 Val Ala Gly Gly Leu Trp Trp Pro Tyr Arg Ile Glu Pro Val Ala Leu
 45 50 55

35

gcc cag gcc tgg gcg ctg cgt tcc ctg gac gtg tac gag gag ctg gcg 246
 Ala Gln Ala Trp Ala Leu Arg Ser Leu Asp Val Tyr Glu Glu Leu Ala
 60 65 70

40

gca cgg ccc ggg cag acc ggc gta cgc atg ctc gaa ggg gtg ctc ggc 294
 Ala Arg Pro Gly Gln Thr Gly Val Arg Met Leu Glu Gly Val Leu Gly
 75 80 85

gag acc ggc ctg gac gag gtg gac ggg tgg gcc gcg gcc cgg ctg ccg 342
 Glu Thr Gly Leu Asp Glu Val Asp Gly Trp Ala Ala Ala Arg Leu Pro
 90 95 100

45

ggg ctg cgc gcg gcg agc gcc gcc gag tac gcc ggg acg ggg ctg tgg 390

	Gly	Leu	Arg	Ala	Ala	Ser	Ala	Ala	Glu	Tyr	Ala	Gly	Thr	Gly	Leu	Trp	
	105					110					115					120	
5	gcg	cgg	ctg	ccg	ctc	atc	gac	atg	tcg	acc	cat	ctg	ccg	tgg	ctg	cgg	438
	Ala	Arg	Leu	Pro	Leu	Ile	Asp	Met	Ser	Thr	His	Leu	Pro	Trp	Leu	Arg	
					125					130					135		
10	gag	cgg	ctg	ctg	gcc	gcg	ggc	ggc	acg	gtg	gag	gac	cgc	gcg	gtg	acc	486
	Glu	Arg	Leu	Leu	Ala	Ala	Gly	Gly	Thr	Val	Glu	Asp	Arg	Ala	Val	Thr	
					140					145					150		
15	gat	ctg	gcc	gag	gcg	gac	gcg	ccg	gtg	gtg	gtc	aac	tgc	acc	ggc	ctg	534
	Asp	Leu	Ala	Glu	Ala	Asp	Ala	Pro	Val	Val	Val	Asn	Cys	Thr	Gly	Leu	
			155					160					165				
	ggc	gcc	cgg	gag	ctg	gtg	ccg	gac	ccg	gcg	gta	cgg	ccg	gtg	cgc	gga	582
	Gly	Ala	Arg	Glu	Leu	Val	Pro	Asp	Pro	Ala	Val	Arg	Pro	Val	Arg	Gly	
		170						175					180				
20	cag	ctg	gtc	gtc	gtg	gag	aac	ccc	ggc	atc	cac	aac	tgg	ctg	gtc	gcg	630
	Gln	Leu	Val	Val	Val	Glu	Asn	Pro	Gly	Ile	His	Asn	Trp	Leu	Val	Ala	
	185						190					195				200	
25	gcc	gac	gcg	gac	tcc	ggg	gag	acg	acg	tac	ttc	ctt	ccg	cag	ccg	gga	678
	Ala	Asp	Ala	Asp	Ser	Gly	Glu	Thr	Thr	Tyr	Phe	Leu	Pro	Gln	Pro	Gly	
					205					210					215		
30	cgg	ctc	ctg	ctg	ggc	ggc	acg	gct	gag	gag	gac	gcc	tgg	tcg	acc	gag	726
	Arg	Leu	Leu	Leu	Gly	Gly	Thr	Ala	Glu	Glu	Asp	Ala	Trp	Ser	Thr	Glu	
				220					225					230			
35	ccg	gac	ccg	gag	gtc	gcg	gcg	gcc	atc	gtg	cga	cgg	tgc	gcg	gcc	ctg	774
	Pro	Asp	Pro	Glu	Val	Ala	Ala	Ala	Ile	Val	Arg	Arg	Cys	Ala	Ala	Leu	
			235					240					245				
	cgt	ccc	gag	atc	gcc	gga	gcg	cgg	gtg	ctc	gcg	cac	ctg	gtg	ggg	ctg	822
	Arg	Pro	Glu	Ile	Ala	Gly	Ala	Arg	Val	Leu	Ala	His	Leu	Val	Gly	Leu	
		250					255					260					
40	cgg	ccg	gcc	cgg	gac	gcg	gtc	cgg	ctg	gag	cgc	ggg	acg	ctg	ccg	gac	870
	Arg	Pro	Ala	Arg	Asp	Ala	Val	Arg	Leu	Glu	Arg	Gly	Thr	Leu	Pro	Asp	
	265					270					275					280	
45	ggg	cgc	cgg	ctg	gtg	cac	aac	tac	ggt	cac	ggc	ggc	gcg	ggc	gtc	acc	918
	Gly	Arg	Arg	Leu	Val	His	Asn	Tyr	Gly	His	Gly	Gly	Ala	Gly	Val	Thr	

	285	290	295	
	gtg gcc tgg ggc tgc gct cag gag gcg gcc cgg ctc gcc tcctga			963
	Val Ala Trp Gly Cys Ala Gln Glu Ala Ala Arg Leu Ala			
5	300	305		
	<210> 12			
	<211> 309			
10	<212> PRT			
	<213> Streptomyces coelicolor			
	<220>			
	<221> misc_feature			
15	<222> (880)..(936)			
	<223> DAAO signature			
	<400> 12			
20	Gly Glu Val Val Val Val Gly Gly Gly Val Ile Gly Leu Thr Thr Ala			
	1 5 10 15			
	Val Val Leu Ala Glu Arg Gly Arg Arg Val Arg Leu Trp Thr Arg Glu			
	20 25 30			
25	Pro Ala Glu Arg Thr Thr Ser Val Val Ala Gly Gly Leu Trp Trp Pro			
	35 40 45			
	Tyr Arg Ile Glu Pro Val Ala Leu Ala Gln Ala Trp Ala Leu Arg Ser			
	50 55 60			
30	Leu Asp Val Tyr Glu Glu Leu Ala Ala Arg Pro Gly Gln Thr Gly Val			
	65 70 75 80			
	Arg Met Leu Glu Gly Val Leu Gly Glu Thr Gly Leu Asp Glu Val Asp			
35	85 90 95			
	Gly Trp Ala Ala Ala Arg Leu Pro Gly Leu Arg Ala Ala Ser Ala Ala			
	100 105 110			
40	Glu Tyr Ala Gly Thr Gly Leu Trp Ala Arg Leu Pro Leu Ile Asp Met			
	115 120 125			
	Ser Thr His Leu Pro Trp Leu Arg Glu Arg Leu Leu Ala Ala Gly Gly			
45	130 135 140			

Thr Val Glu Asp Arg Ala Val Thr Asp Leu Ala Glu Ala Asp Ala Pro
 145 150 155 160
 Val Val Val Asn Cys Thr Gly Leu Gly Ala Arg Glu Leu Val Pro Asp
 5 165 170 175
 Pro Ala Val Arg Pro Val Arg Gly Gln Leu Val Val Val Glu Asn Pro
 180 185 190
 Gly Ile His Asn Trp Leu Val Ala Ala Asp Ala Asp Ser Gly Glu Thr
 10 195 200 205
 Thr Tyr Phe Leu Pro Gln Pro Gly Arg Leu Leu Leu Gly Gly Thr Ala
 15 210 215 220
 Glu Glu Asp Ala Trp Ser Thr Glu Pro Asp Pro Glu Val Ala Ala Ala
 225 230 235 240
 Ile Val Arg Arg Cys Ala Ala Leu Arg Pro Glu Ile Ala Gly Ala Arg
 20 245 250 255
 Val Leu Ala His Leu Val Gly Leu Arg Pro Ala Arg Asp Ala Val Arg
 260 265 270
 Leu Glu Arg Gly Thr Leu Pro Asp Gly Arg Arg Leu Val His Asn Tyr
 25 275 280 285
 Gly His Gly Gly Ala Gly Val Thr Val Ala Trp Gly Cys Ala Gln Glu
 290 295 300
 30
 Ala Ala Arg Leu Ala
 305
 <210> 13
 35 <211> 1038
 <212> DNA
 <213> Candida boidinii
 <220>
 40 <221> CDS
 <222> (1)..(1035)
 <223> coding for DAAO
 45

	<400> 13																	
	atg	ggt	gat	caa	att	gtt	gtt	ctt	ggt	tcc	ggt	att	att	ggt	tta	tat	48	
	Met	Gly	Asp	Gln	Ile	Val	Val	Leu	Gly	Ser	Gly	Ile	Ile	Gly	Leu	Tyr		
	1				5					10					15			
5																		
	act	aca	tac	tgt	tta	atc	tat	gag	gct	gga	tgt	gct	cca	gct	aaa	att	96	
	Thr	Thr	Tyr	Cys	Leu	Ile	Tyr	Glu	Ala	Gly	Cys	Ala	Pro	Ala	Lys	Ile		
				20					25					30				
10																		
	act	att	gtt	gct	gaa	ttt	tta	cca	ggt	gat	caa	tct	aca	tta	tat	aca	144	
	Thr	Ile	Val	Ala	Glu	Phe	Leu	Pro	Gly	Asp	Gln	Ser	Thr	Leu	Tyr	Thr		
			35					40						45				
	tct	cca	tgg	gca	ggt	ggt	aat	ttt	tct	tgt	att	tca	cca	gct	gat	gat	192	
15	Ser	Pro	Trp	Ala	Gly	Gly	Asn	Phe	Ser	Cys	Ile	Ser	Pro	Ala	Asp	Asp		
		50					55					60						
	aca	aca	ttg	gct	tat	gat	aaa	ttc	aca	tat	ctt	aat	tta	ttc	aag	att	240	
20	Thr	Thr	Leu	Ala	Tyr	Asp	Lys	Phe	Thr	Tyr	Leu	Asn	Leu	Phe	Lys	Ile		
	65					70					75				80			
	cac	aaa	aaa	tta	ggt	gga	cca	gaa	tgt	gga	tta	gat	aat	aag	cca	agt	288	
	His	Lys	Lys	Leu	Gly	Gly	Pro	Glu	Cys	Gly	Leu	Asp	Asn	Lys	Pro	Ser		
				85						90					95			
25																		
	act	gaa	tat	tgg	gat	ttt	tat	cct	ggt	gat	gaa	aaa	gtc	aat	tct	tta	336	
	Thr	Glu	Tyr	Trp	Asp	Phe	Tyr	Pro	Gly	Asp	Glu	Lys	Val	Asn	Ser	Leu		
				100					105					110				
30																		
	aaa	caa	tat	ctt	aaa	gat	ttt	aaa	gtt	att	cca	aaa	tca	gaa	tta	cca	384	
	Lys	Gln	Tyr	Leu	Lys	Asp	Phe	Lys	Val	Ile	Pro	Lys	Ser	Glu	Leu	Pro		
			115					120					125					
	gaa	ggt	gtt	gaa	tat	ggt	att	agt	tat	act	aca	tgg	aat	ttc	aac	tgt	432	
35	Glu	Gly	Val	Glu	Tyr	Gly	Ile	Ser	Tyr	Thr	Thr	Trp	Asn	Phe	Asn	Cys		
		130					135					140						
	cct	gtt	ttc	tta	caa	aat	atg	gct	aat	ttt	tta	aat	aaa	aga	aat	gtt	480	
	Pro	Val	Phe	Leu	Gln	Asn	Met	Ala	Asn	Phe	Leu	Asn	Lys	Arg	Asn	Val		
40						150					155					160		
	acc	att	att	aga	aaa	cat	tta	aca	cat	att	tct	caa	gct	tat	tta	aca	528	
	Thr	Ile	Ile	Arg	Lys	His	Leu	Thr	His	Ile	Ser	Gln	Ala	Tyr	Leu	Thr		
					165					170					175			
45																		

	gtt aat aca aaa gtt gtt ttc aac tgt aca ggt att ggt gct gct gat	576
	Val Asn Thr Lys Val Val Phe Asn Cys Thr Gly Ile Gly Ala Ala Asp	
	180 185 190	
5	tta ggt ggt gtt aaa gat gaa aaa gtt tat cca act aga gga caa gtt	624
	Leu Gly Gly Val Lys Asp Glu Lys Val Tyr Pro Thr Arg Gly Gln Val	
	195 200 205	
10	gtt gtt gtt aga gct cca cat att caa gaa aat aaa atg aga tgg ggt	672
	Val Val Val Arg Ala Pro His Ile Gln Glu Asn Lys Met Arg Trp Gly	
	210 215 220	
15	aaa gac tat gct act tat att att cca aga cca tat tct aat ggt gaa	720
	Lys Asp Tyr Ala Thr Tyr Ile Ile Pro Arg Pro Tyr Ser Asn Gly Glu	
	225 230 235 240	
20	tta gtc tta ggt ggt ttc tta caa aag gat aat tgg aca ggt aat act	768
	Leu Val Leu Gly Gly Phe Leu Gln Lys Asp Asn Trp Thr Gly Asn Thr	
	245 250 255	
25	ttt ggt ttt gaa act gat gat att gtt agt aga act aca tct tta tta	816
	Phe Gly Phe Glu Thr Asp Asp Ile Val Ser Arg Thr Thr Ser Leu Leu	
	260 265 270	
30	cca aag att tta gat gaa cca ctt cat att att aga gtt gca gct ggt	864
	Pro Lys Ile Leu Asp Glu Pro Leu His Ile Ile Arg Val Ala Ala Gly	
	275 280 285	
35	tta aga cca agt aga cat ggt ggt cca aga att gaa gct gaa gtt tgt	912
	Leu Arg Pro Ser Arg His Gly Gly Pro Arg Ile Glu Ala Glu Val Cys	
	290 295 300	
40	gat aac caa aaa gtt aaa gct aaa ctt tag	1038
	Asp Asn Gln Lys Val Lys Ala Lys Leu	
	340 345	

<210> 14
 <211> 345
 <212> PRT
 <213> Candida boidinii

5

<400> 14
 Met Gly Asp Gln Ile Val Val Leu Gly Ser Gly Ile Ile Gly Leu Tyr
 1 5 10 15

10 Thr Thr Tyr Cys Leu Ile Tyr Glu Ala Gly Cys Ala Pro Ala Lys Ile
 20 25 30

Thr Ile Val Ala Glu Phe Leu Pro Gly Asp Gln Ser Thr Leu Tyr Thr
 35 40 45

15

Ser Pro Trp Ala Gly Gly Asn Phe Ser Cys Ile Ser Pro Ala Asp Asp
 50 55 60

20 Thr Thr Leu Ala Tyr Asp Lys Phe Thr Tyr Leu Asn Leu Phe Lys Ile
 65 70 75 80

His Lys Lys Leu Gly Gly Pro Glu Cys Gly Leu Asp Asn Lys Pro Ser
 85 90 95

25 Thr Glu Tyr Trp Asp Phe Tyr Pro Gly Asp Glu Lys Val Asn Ser Leu
 100 105 110

Lys Gln Tyr Leu Lys Asp Phe Lys Val Ile Pro Lys Ser Glu Leu Pro
 115 120 125

30

Glu Gly Val Glu Tyr Gly Ile Ser Tyr Thr Thr Trp Asn Phe Asn Cys
 130 135 140

35 Pro Val Phe Leu Gln Asn Met Ala Asn Phe Leu Asn Lys Arg Asn Val
 145 150 155 160

Thr Ile Ile Arg Lys His Leu Thr His Ile Ser Gln Ala Tyr Leu Thr
 165 170 175

40 Val Asn Thr Lys Val Val Phe Asn Cys Thr Gly Ile Gly Ala Ala Asp
 180 185 190

Leu Gly Gly Val Lys Asp Glu Lys Val Tyr Pro Thr Arg Gly Gln Val
 195 200 205

45

Val Val Val Arg Ala Pro His Ile Gln Glu Asn Lys Met Arg Trp Gly
 210 215 220

5 Lys Asp Tyr Ala Thr Tyr Ile Ile Pro Arg Pro Tyr Ser Asn Gly Glu
 225 230 235 240

10 Leu Val Leu Gly Gly Phe Leu Gln Lys Asp Asn Trp Thr Gly Asn Thr
 245 250 255

Phe Gly Phe Glu Thr Asp Asp Ile Val Ser Arg Thr Thr Ser Leu Leu
 260 265 270

15 Pro Lys Ile Leu Asp Glu Pro Leu His Ile Ile Arg Val Ala Ala Gly
 275 280 285

Leu Arg Pro Ser Arg His Gly Gly Pro Arg Ile Glu Ala Glu Val Cys
 290 295 300

20 Glu Glu Gly Lys Leu Thr Ile His Asn Tyr Gly Ala Ser Gly Tyr Gly
 305 310 315 320

Tyr Gln Ala Gly Tyr Gly Met Ser Tyr Glu Ala Val Lys Leu Leu Val
 325 330 335

25 Asp Asn Gln Lys Val Lys Ala Lys Leu
 340 345

30 <210> 15
 <211> 12466
 <212> DNA
 <213> vector daaoSceITetON

35 <220>
 <221> misc_feature
 <222> (38)..(183)
 <223> Agrobacterium right border

40 <220>
 <221> misc_feature
 <222> (445)..(462)
 <223> recognition / cleavage site for I-SceI endonuclease

<220>
<221> terminator
<222> (196)..(400)
<223> complementary: 35S terminator
5
<220>
<221> misc_feature
<222> (515)..(1222)
<223> complementary: coding for I-SceI endonuclease
10
<220>
<221> promoter
<222> (1270)..(1660)
<223> complementary: coding for pTOP10P tetracyclin regulatable promoter
15
<220>
<221> terminator
<222> (1735)..(1990)
<223> complementary: NOS terminator
20
<220>
<221> misc_feature
<222> (2067)..(3173)
<223> complementary: coding for Rhodotorula gracilis D-amino acid oxidase
25
<220>
<221> promoter
<222> (3217)..(5028)
30 <223> complementary: Arabidopsis thaliana nitrilase I promoter
<220>
<221> terminator
<222> (5118)..(5343)
35 <223> complementary: OCS terminator
<220>
<221> misc_feature
<222> (5418)..(6425)
40 <223> complementary: coding for tetracyclin repressor rtTA
<220>
<221> promoter
<222> (6479)..(7341)
45 <223> complementary: coding for Pisum sativum ptxA promoter

<220>

<221> terminator

<222> (7345)..(7549)

5 <223> complementary: coding for 23S terminator (functioning as homology sequence)

<220>

<221> misc_feature

<222> (7618)..(7834)

10 <223> Agrobacterium left border

<400> 15

	aatattcaaa	caaacacata	cagcgcgact	tatcatggac	atacaaattg	acgaacggat	60
	aaaccttttc	acgccctttt	aaatatccga	ttattctaata	aaacgctctt	ttctcttagg	120
15	tttaccggcc	aatatatcct	gtcaaacact	gatagtttaa	actgaaggcg	ggaaacgaca	180
	atcagatctg	gtaccgggtc	actggatttt	ggtttttagga	attagaaatt	ttattgatag	240
	aagtatttta	caaatacaaa	tacatactaa	gggttttctta	tatgctcaac	acatgagcga	300
	aacctataa	gaaccctaata	tcccttatct	gggaactact	cacacattat	tctggagaaa	360
	aatagagaga	gatagatttg	tagagagaga	ctggtgattt	ttgcgccggg	taccccaaac	420
20	tgtctcacga	cgttttgaac	ccagattacc	ctgttatccc	tagtcgagcg	gccgccagtg	480
	tgatggatat	ctgcagaatt	cgccctttta	gatcttattt	caggaaagtt	tcggaggaga	540
	tagtggtcgg	cagtttgtac	atcatctgcg	ggatcaggta	cggtttgatc	aggttgtaga	600
	agatcaggta	agacatagaa	tcgatgtaga	tgatcggttt	gtttttgttg	atttttacgt	660
	aacagttcag	ttggaatttg	ttacgcagac	ccttaaccag	gtattctact	tcttcgaaag	720
25	tgaaagactg	ggtgttcagt	acgatcgatt	tgttggtaga	gtttttgttg	taatccatt	780
	taccaccatc	atccatgaac	cagtatgcca	gagacatcgg	ggtcaggtag	ttttcaacca	840
	ggttggtcgg	gatgggtttt	ttgttggtta	cgatgaacag	gctagccagt	ttgttgaaag	900
	cttggtgttt	gaaagtctgg	gcgccccagg	tgattaccag	gttaccagg	tggttaacac	960
	gttctttttt	gtgcggcggg	gacagtacc	actgatcgta	cagcagacat	acgtggtcca	1020
30	tgtatgcttt	gtttttccac	tcgaactgca	tacagtaggt	tttaccttca	tcacgagaac	1080
	ggatgtaagc	atcaccagg	atcagaccga	tacctgcttc	gaactgttcg	atgttcagtt	1140
	cgatcagctg	ggatttgtat	tcttttcagca	gttttagagtt	cggacccagg	ttcattacct	1200
	ggtttttttt	gatgtttttc	atatggtcga	ctaaagggcg	aattccagca	cactggcggc	1260
	cgttactagc	ccgggctcga	gcaaattgtc	agaaaggcct	tatatacgta	aaggggtctt	1320
35	cgaagactag	atcactctat	ctcgagttta	ccactcccta	tcagtgatag	agaaaagtga	1380
	aagtcgagtt	taccactccc	tatcagtgat	agagaaaagt	gaaagtcgag	tttaccactc	1440
	cctatcagtg	atagagaaaa	gtgaaagtcg	agtttaccac	tccctatccg	tgatagagaa	1500
	aagtgaaagt	cgagtttacc	actccctatc	agtgatagag	aaaagtga	gtcgagttta	1560
	ccactcccta	tcagtgatag	agaaaatgaa	agtcgagttt	accactccct	atcagtgata	1620
40	gagaaaagtg	aaagtcgagc	tcggtaccga	gctcgaattc	agcacactgg	cggccgttac	1680
	tagtggatca	attcactggc	cgtcgtttta	caacgactca	gagcttgaca	ggaggcccga	1740
	tctagtaaca	tagatgacac	cgcgcgcgat	aatttatcct	agtttgcgcg	ctatattttg	1800
	ttttctatcg	cgtattaaat	gtataattgc	gggactctaa	tcataaaaac	ccatctcata	1860
	aataacgtca	tgcattacat	gttaattatt	acatgcttaa	cgtaattcaa	cagaaattat	1920
45	atgataatca	tcgcaagacc	ggcaacagga	ttcaatctta	agaaacttta	ttgccaaatg	1980

	tttgaacgat	cggggatcat	ccgggtctgt	ggcgggaact	ccacgaaaat	atccgaacgc	2040
	agcaagatct	agagcttggg	tcccgcctac	aacttcgact	cccgcgccgc	gccgtggtac	2100
	cgctggaacg	cctcgtcgac	gagctgcgcg	acatcctccg	ccgcgccccca	actctgctgg	2160
	tatccccgcac	tcgagaagcc	atacgcatgc	acaagcgtga	cctcctttctc	cttcgccgct	2220
5	cgtgcgctgc	ccctgccgag	cgagaggggc	gactttgtcc	ggtcgagagg	caggacgatc	2280
	cgttctgcct	caacgcgggg	tccgcctcgt	cgtgcaggtc	gcaagccgac	gttgtggcgg	2340
	aggacctcga	tgcttcgat	cgttccgtcg	ctcgagatgg	tcgggtcgag	gcgcaagcag	2400
	tgcttgagga	tcgcctggac	cgtctctggg	ttgacagaca	agtcccagtc	tcccacgccg	2460
	tacgtcccgc	cgcagatgac	ttcgccacct	ggtcggggaa	tgatgtaggc	gggagaagcg	2520
10	gggtcggacg	agtccatcgt	gcatcgcttg	catggggact	tgacgaggac	ggtttgcccg	2580
	cggattggct	cggcggcttg	gtcgtcgatg	cccgcaatcg	acttggcgcc	aagtcccgta	2640
	gcgttgacca	ccaaatccgc	accgtcgaac	gcctgctcaa	gcgacgtaac	ggtcggtctc	2700
	tcaaacgtcg	cgcgcgagctt	ctgcagctct	cttgcaaggt	actggcagta	ctttgggtgcg	2760
	tggacggaga	gggtgtcgta	ggttacgccg	atagcgccag	gtggacattc	ggaagatggg	2820
15	agggggcggt	aatttggcgt	gatgtccttg	taccagtgcc	cgagcaagcc	gtcttcgttc	2880
	tgcgcgaaac	gcctcgtccc	cttgagccac	atggcatggc	ccgtcgggac	caactcgacc	2940
	cacttcttga	aagtcgattc	ttcccatttt	gcttgctcgag	gaccgtctgt	aagcgtcatg	3000
	aaaggcgtcc	aattcgcgcc	agcccatggt	gaagcgaaag	tctggctcga	gacgtcctcc	3060
	ggcaagtcgc	gcgcgagaat	atgcacgctg	tagcccttcc	gagcgaggat	gagggcgctg	3120
20	ctcagaccga	taacgcctga	tccgaggaca	acgacgcgct	tctgcgagtg	catgggcccct	3180
	cgactagagt	cgagatccga	tatcgcccgg	gctcgagtct	ttgttttttta	ctttgggttca	3240
	tgacactcag	agacttgaga	gaagcaatat	atagactttt	ttttgttttt	tttttgtggt	3300
	cacgtttatt	ttcctattgg	agacggtaac	gaagatcgaa	cctgtggtgg	aaatgaaaca	3360
	aggtgggact	agcccacgtg	gtttcttttc	tctgcattga	tttgtttttg	ttttttttgt	3420
25	aaagttcaca	tcaaacctac	taataattga	gaagaaaaat	aaaatctatt	gattgattaa	3480
	accagccgat	gctttatgtc	tgaatataaa	aaagaagtga	aaaccccggt	taagaattac	3540
	aacggtgggt	tacaaagtat	ttggacacaa	taaatccaaa	cgaaataaaa	caaaatggag	3600
	aactaccaa	taaaaaaca	ataaaaaact	taaaagaatt	tattccattt	tttttcccgt	3660
	agaatttatt	cttttatgga	ttccttaaat	ccatatattga	tgcatatttg	ttcctcataa	3720
30	taggtaataa	tatatactat	gttatagata	tgtttcta	tcgtattaac	ctaccttttt	3780
	ttggtcgtac	gattctacct	aataatattg	aacggaattg	atgttttgga	ccacttagaa	3840
	agtatttttt	ttttggtttg	tcttagctgt	atttcattaa	atataaat	aaataagaaa	3900
	tgtcataaat	aaaatttgac	gtatagattt	tttaaatcca	ttttatgtta	tttaatat	3960
	gaaatgtgag	tttggctcct	atttaatcct	aggatggggt	aataactaagt	tttcctta	4020
35	gaattatctc	agagaaactg	gattaaataa	actaaaaaat	agatcaatgt	gttttggtcc	4080
	ggtcaaatat	ctttggattt	actattattg	gcgaaaagaa	agtctcatat	agtaaatcat	4140
	attcctacaa	gagaaatcaa	aatttttgaa	ttaacatgga	ttgtatagtt	tcttatataa	4200
	ccaattagtt	cgcataca	aaaccaaacc	ccaattaata	atcaaacggg	cttggttagga	4260
	atatttcatt	gcagctttca	gataaaagaa	aaaaacacac	actcaagtct	tttatttc	4320
40	ctttcttact	tgacaggaact	caaattccac	tttgccactt	ttcttttaca	ataaacacaa	4380
	attgtcaatg	aaacgaaata	gtctttttat	gcaaacactg	tttgtctttt	ttcgatcacg	4440
	tttctgattg	tgacagccat	ccatatatat	agggaatgta	aaacaacaac	atgtgaagtc	4500
	acataacgt	aatggtttag	catagcttct	attttcgttg	tcaatattag	tcattccaaa	4560
	acattttttaa	gaaaaataaa	ttaatatatg	tatatctctg	gaactaatgt	atgtggaaat	4620
45	acagtaactt	aattattaaa	cattctaaat	gcaaatatgc	aaagaaaaaa	aagaaaagaa	4680

	cacaactgaa	atcaaagcca	gattcataat	aattggctac	atgggttgtag	aatgtagggt	4740
	aacacaacat	ccagaattga	acactcaaat	tggatgatag	atggataatc	tttagataca	4800
	agagaattgg	ttctcttcca	ttattaacga	aaataaagaa	aaaaagttta	gcataaaagt	4860
	ttgaaactca	acataacatt	ttgaacttga	ctccttcata	ggagtgcacat	gaactgacga	4920
5	atcacaaccg	attacttgtt	tgagtcattt	tccgctttct	ccaccttcga	aatgaatgtg	4980
	accggtttct	tcgggtgctc	atttacggtc	aagtgtaaaa	catctgggtc	cgagggtacct	5040
	ggtagggata	acagggtaat	ctgggttcaa	aacgtcgtga	gacagtttgg	tgcagggtcga	5100
	aattcgagct	cggtaccaat	tcccatcttg	aaagaaatat	agtttaataa	tttattgata	5160
	aaataacaag	tcagggtatta	tagtccaagc	aaaaacataa	atttattgat	gcaagtttaa	5220
10	attcagaaat	atttcaataa	ctgattatat	cagctggtag	attgccgtag	atgaaagact	5280
	gagtgcgata	ttatgtgtaa	tacataaatt	gatgatatag	ctagcttagc	tcacgggggg	5340
	atcttgcgcc	gggtaccgag	ctcggtagca	attcccgagg	ctgtagccga	cgatgggtgcg	5400
	ccaggagagt	tgttgatcta	cccaccgtac	tcgtcaattc	caagggcatc	ggtaaacatc	5460
	tgctcaaact	cgaagtcggc	catatccaga	gcgccgtagg	gggcggagtc	gtggggggta	5520
15	aatcccgac	ccggggaatc	cccgcccccc	aacatgtcca	gatcgaaatc	gtctagcgcg	5580
	tcggcatgcg	ccatcgccac	gtcctcgccg	tctaagtga	gctcgtcccc	caggctgaca	5640
	tcggtcgggg	gggccgtcga	cagtctgcgc	gtgtgtcccc	cggggagaaa	ggacaggcgc	5700
	ggagccgcca	gccccgcctc	ttcggggggc	tcgtcgtccg	ggagatcgag	caggccctcg	5760
	atggtagacc	cgtaattgtt	tttcgtacgc	gcgcggctgt	acgcggaccc	actttcacat	5820
20	ttaagtgtgt	tttctaattc	gcataatgat	aattcaaggc	cgaataagaa	ggctggctct	5880
	gcaccttggt	gatcaaataa	ttcgatagct	tgctcgtata	atggcggcat	actatcagta	5940
	gtaggtgttt	ccctttcttc	tttagcgact	tgatgctctt	gatcttccaa	tacgcaacct	6000
	aaagtaaaat	gccccacagc	gctgagtga	tataacgcgt	tctctagtga	aaaaccttgt	6060
	tggcataaaa	aggctaattg	atttttcgaga	gtttcatact	gtttttctgt	aggccgtgta	6120
25	tctgaatgta	cttttgctcc	attgcgatga	cttagtaaag	cacatctaaa	acttttagcg	6180
	ttattgcgta	aaaaatcttg	ccagcttttc	ccttttaaa	ggcaaaagt	agtatgggtc	6240
	ctatctaaca	tctcaatggc	taaggcgctc	agcaaagccc	gcttattttt	tacatgccaa	6300
	tacagtgtag	gctgctctac	accaagcttc	tgggcgagtt	tacgggttgt	taaaccttcg	6360
	attccgacct	cattaagcag	ctctaattgc	ctgttaatac	ctttactttt	atctaattct	6420
30	gacatggctc	atcgactcta	gactagtga	tccgatatac	cccgggctcg	actctagagt	6480
	ttcgaagatt	ttagtgtaat	gtgtgtgctc	actactatga	agctttgcac	ttaaaaaat	6540
	agaatgagt	atgaggttta	tatggtgaaa	aaaactatga	aattttgata	ttttgatata	6600
	tctttctcgt	gagtcattat	cacggaccat	gttgcagcaa	attggaatta	aactattcat	6660
	tttttatgtt	aaatcattga	ttgattttta	gtgggcctcg	ttacatatct	aagagttaga	6720
35	atgaattcaa	acaaactagg	ccagaaaaaa	ggatgtgggg	ccattttttt	gtgtcttaag	6780
	aatttggtta	tttttttcat	ggataagggg	aatcaatgga	aaaagtttga	tgtactagag	6840
	gacatttttt	taacatgtag	tgacaagtag	tgctattatt	cgaccctgta	tgaaaggggc	6900
	aatcttaatc	tttttttcat	aaatctgcac	atgtgatgct	ttaattatgc	tttagacttt	6960
	gtgctaaact	attggtaatt	tcttttttga	atcgaatcaa	gtatctttta	aactatgtat	7020
40	gaaatgtgtc	atcctaataa	caacattttg	ctagttttag	actttgatgt	ttatatgctt	7080
	aatggaagaa	gcaatatgtt	gatgtttatt	gggtaaaaga	aagggacttg	attgagtatg	7140
	taattgacaa	ctatgatttt	atattggatt	tgatattcct	aacattaatt	taagtgtgtg	7200
	ggtttcaaag	catgttatgc	tagtgattct	tgtgtttgat	gcttgaaaaa	tctacattca	7260
	tccttgaatg	gagggacaaa	ctttgaatga	cttttgaata	ggtgtaaaaa	ccaatcctcc	7320
45	ctcagcttca	caaaaaattg	cggacgggtc	ctggattttg	gttttaggaa	ttagaaattt	7380

	tattgataga	agtatttttac	aaatacaaat	acatactaag	ggttttcttat	atgctcaaca	7440
	catgagcgaa	accctataag	aaccctaatt	cccttatctg	ggaactactc	acacattatt	7500
	ctggagaaaa	atagagagag	atagattttgt	agagagagac	tggtgatttt	ccggggggtc	7560
	ctctagagtc	gaggtaccga	gctcgaattc	actggccgtc	gtttttacaac	gactcagtac	7620
5	tgcttggttaa	taattgtcat	tagattgttt	ttatgcatag	atgcactcga	aatcagccaa	7680
	tttttagacaa	gtatcaaacg	gatgttaatt	cagtacatta	aagacgtccg	caatgtgtta	7740
	ttaagttgtc	taagcgtcaa	tttgttttaca	ccacaatata	tcctgccacc	agccagccaa	7800
	cagctccccg	accggcagct	cggcacaaaa	tcaccacgcg	tctaaaaagg	tgatgtgtat	7860
	ttgagtaaaa	cagcttgctg	catgcggtcg	ctgcgtatat	gatgcgatga	gtaaataaac	7920
10	aaatacgcga	ggggaacgca	tgaaggttat	cgctgtactt	aaccagaaag	gcgggtcagg	7980
	caagacgacc	atcgcaaccc	atctagcccc	cgccttgcaa	ctcgccgggg	ccgatgttct	8040
	gttagtcgat	tccgatcccc	agggcagtg	ccgcgattgg	gcggccgtgc	gggaagatca	8100
	accgctaacc	gttgctcgga	tcgaccgccc	gacgattgac	cgcgacgtga	aggccatcgg	8160
	ccggcgcgac	ttcgtagtga	tcgacggagc	gccccaggcg	gcggacttgg	ctgtgtccgc	8220
15	gatcaaggca	gccgacttcg	tgctgattcc	ggtgcagcca	agcccttacg	acatatgggc	8280
	caccgccgac	ctggtggagc	tggttaagca	gcgcattgag	gtcacggatg	gaaggctaca	8340
	agcggccttt	gtcgtgtcgc	gggcgatcaa	aggcacgcgc	atcggcggtg	aggttgccga	8400
	ggcgctggcc	gggtacgagc	tgcccattct	tgagtcccgt	atcacgcagc	gcgtgagcta	8460
	cccaggcact	gccgccgccc	gcacaaccgt	tcttgaatca	gaaccgcagg	gcgacgctgc	8520
20	ccgcgaggtc	caggcgctgg	ccgctgaaat	taaatcaaaa	ctcatttgag	ttaatgaggt	8580
	aaagagaaaa	tgagcaaaaag	cacaaacacg	ctaagtgcgc	gccgtccgag	cgcacgcagc	8640
	agcaaggctg	caacgttggc	cagcctggca	gacacgccag	ccatgaagcg	ggtcaacttt	8700
	cagttgccgg	cggaggatca	caccaagctg	aagatgtacg	cggtagccca	aggcaagacc	8760
	attaccgagc	tgctatctga	atacatcgcg	cagctaccag	agtaaatagag	caaatagaata	8820
25	aatgagtaga	tgaatttttag	cggctaaagg	aggcggcatg	gaaaatcaag	aacaaccagg	8880
	caccgacgcc	gtggaatgcc	ccatgtgtgg	aggaaacggc	ggttgggccag	gcgtaagcgg	8940
	ctgggttgct	tgccggccct	gcaatggcac	tggaaccccc	aagcccagag	aatcggcgctg	9000
	agcggtcgca	aaccatccgg	cccgggtacaa	atcggcgcg	cgctgggtga	tgacctggtg	9060
	gagaagttga	aggccgcgca	ggccgcccag	cggcaacgca	tcgaggcaga	agacgccccg	9120
30	gtgaatcgtg	gcaagggggc	gctgatcgaa	tccgcaaaga	atcccggcaa	ccgccggcag	9180
	ccggtgcgcc	gtcgattagg	aagccgcccc	agggcgacga	gcaaccagat	tttttcgctc	9240
	cgatgctcta	tgacgtgggc	acccgcgata	gtcgcagcat	catggacgtg	gccgttttcc	9300
	gtctgtcgaa	gcgtgaccga	cgagctggcg	agggtgatccg	ctacgagctt	ccagacgggc	9360
	acgtagaggt	ttccgcaggg	ccggccggca	tggccagtgt	gtgggattac	gacctggtac	9420
35	tgatggcggt	ttcccatcta	accgaatcca	tgaaccgata	ccgggaagg	aaggagagaca	9480
	agcccggccg	cgtgttccgt	ccacacgttg	cggacgtact	caagtctctg	cggcgagccg	9540
	atggcgga	gcagaaagac	gacctggtag	aaacctgcat	tcggttaaac	accacgcacg	9600
	ttgccatgca	gcgtacgaag	aaggccaaga	acggccgcct	ggtgacggta	tccgaggggtg	9660
	aagccttgat	tagccgctac	aagatcgtaa	agagcgaaac	cgggcggccg	gagtacatcg	9720
40	agatcgagct	agctgattgg	atgtaccgcg	agatcacaga	aggcaagaac	ccggacgtgc	9780
	tgacggttca	ccccgattac	ttttttgatcg	atcccggcat	cggccgtttt	ctctaccgcc	9840
	tggcacgccg	cgccgcaggc	aaggcagaag	ccagatggtt	gttcaagacg	atctacgaac	9900
	gcagtggcag	cgccggagag	ttcaagaagt	tctgtttcac	cgtgcgcaag	ctgatcgggt	9960
	caaatgacct	gccggagtac	gattttgaagg	aggaggcg	gcaggctggc	ccgatccctag	10020
45	tcatgcgcta	ccgcaacctg	atcgaggggc	aagcatccgc	cggttcctaa	tgtacggagc	10080

	agatgctagg	gcaaattgcc	ctagcagggg	aaaaagggtcg	aaaagggtctc	tttcctgtgg	10140
	atagcacgta	cattgggaac	ccaaagccgt	acattgggaa	ccggaacccg	tacattggga	10200
	acccaaagcc	gtacattggg	aaccggtcac	acatgtaagt	gactgatata	aaagagaaaa	10260
	aaggcgattt	ttccgcctaa	aactctttta	aacttattaa	aactcttaaa	accgccttg	10320
5	cctgtgcata	actgtctggc	cagcgcacag	ccgaagagct	gcaaaaagcg	cctacccttc	10380
	ggtcgctgcg	ctccctacgc	cccgccgctt	cgcgtcggcc	tatcgcgggc	tatgcggtgt	10440
	gaaataccgc	acagatgcgt	aaggagaaaa	taccgcatca	ggcgctcttc	cgcttcctcg	10500
	ctcactgact	cgctgcgctc	ggtcgttcgg	ctgcggcgag	cggtatcagc	tcactcaaag	10560
	gcggtaatac	ggttatccac	agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	10620
10	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcgttgc	tggcggtttt	ccataggctc	10680
	cgccccoctg	acgagcatca	caaaaatcga	cgctcaagtc	agagggtggcg	aaacccgaca	10740
	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	tcgtgcgctc	tcctgttccg	10800
	accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	cgggaagcgt	ggcgctttct	10860
	catagctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	ttcgctccaa	gctgggctgt	10920
15	gtgcacgaac	ccccggttca	gcccgaaccg	tgcgccttat	ccggtaacta	tcgtcttgag	10980
	tccaacccgg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	caggattagc	11040
	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	ggtaggctaa	ctacggctac	11100
	actagaagga	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	cggaaaaaga	11160
	gttggttagct	cttgatccgg	caaacaacc	accgctggta	gcggtgggtt	ttttgtttgc	11220
20	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	cttttctacg	11280
	gggtctgacg	ctcagtggaa	cgaaaactca	cgttaaggga	ttttgggtcat	gcatgatata	11340
	tctcccaatt	tgtgtagggc	ttattatgca	cgcttaaaaa	taataaaaagc	agacttgacc	11400
	tgatagtttg	gctgtgagca	attatgtgct	tagtgcatct	aacgcttgag	ttaagccgcg	11460
	ccgcgaagcg	gcgtcggctt	gaacgaattt	ctagctagac	attatttgcc	gactaccttg	11520
25	gtgatctcgc	ctttcacgta	gtggacaaat	tcttccaact	gatctgcgcg	cgaggccaag	11580
	cgatctttctt	cttgtccaag	ataagcctgt	ctagcttcaa	gtatgacggg	ctgatactgg	11640
	gccggcaggc	gctccattgc	ccagtcggca	gcgacatcct	tcggcgcgat	tttgccgggtt	11700
	actgcgctgt	accaaattgcg	ggacaacgta	agcactacat	ttcgctcatc	gccagcccag	11760
	tcggggcggcg	agttccatag	cggttaagggtt	tcatttagcg	cctcaaataag	atcctgttca	11820
30	ggaaccggat	caaagagttc	ctccgccgct	ggacctacca	aggcaacgct	atgttctctt	11880
	gctttttgtca	gcaagatagc	cagatcaatg	tcgatcgtgg	ctggctcgaa	gatacctgca	11940
	agaatgtcat	tgcgctgcca	ttctccaaat	tgcagttcgc	gcttagctgg	ataacgccac	12000
	ggaatgatgt	cgtcgtgcac	aacaatggtg	acttctacag	cgcggagaat	ctcgctctct	12060
	ccagggggaag	ccgaagtttc	caaaagggtcg	ttgatcaaag	ctcgccgcgt	tgtttcatca	12120
35	agccttacgg	tcaccgtaac	cagcaaatca	atatcactgt	gtggcttcag	gccgccatcc	12180
	actgcggagc	cgtacaaatg	tacggccagc	aacgtcgggtt	cgagatggcg	ctcgatgacg	12240
	ccaactacct	ctgatagttg	agtcgatact	tcggcgatca	ccgcttcccc	catgatgttt	12300
	aactttgttt	tagggcgact	gccctgctgc	gtaacatcgt	tgctgctcca	taacatcaaa	12360
	catcgacca	cggcgtaacg	cgcttgctgc	ttggatgcc	gaggcataga	ctgtaccca	12420
40	aaaaaacagt	cataacaagc	catgaaaacc	gccactgcgt	tccatg		12466

<210> 16

<211> 12539

45 <212> DNA

<213> daaoNit-PRecombination

<220>

5

<221> misc_feature

<222> (38)..(183)

<223> Agrobacterium right border

<220>

10

<221> terminator

<222> (196)..(400)

<223> complementary: 35S terminator

<220>

15

<221> misc_feature

<222> (445)..(462)

<223> cleavage / recognition site for I-SceI endonuclease

<220>

20

<221> terminator

<222> (589)..(844)

<223> complementary: nos terminator

<220>

25

<221> misc_feature

<222> (921)..(2027)

<223> complementary: coding for Rhodotorula gracilis D-amino acid oxidase

30

<220>

<221> promoter

<222> (2071)..(3882)

<223> complementary: A-thaliana nitrilase I promoter

35

<220>

<221> terminator

<222> (3972)..(4176)

<223> complementary: 35S terminator

40

<220>

<221> misc_feature

<222> (4251)..(6248)

<223> complementary: coding for beta-glucuronidase

<220>
 <221> promoter
 <222> (6302)..(7619)
 <223> complementar. coding for sTPT promoter

5

<220>
 <221> misc_feature
 <222> (7691)..(7907)
 <223> Agrobacterium left border

10

<400> 16
 aatattcaaa caaacacata cagcgcgact tatcatggac atacaaatgg acgaacggat 60
 aaaccttttc acgccctttt aaatatccga ttatttctaat aaacgctctt ttctcttagg 120
 tttaccgcgc aatatatcct gtcaaacact gatagtttta actgaaggcg ggaaacgaca 180
 15 atcagatctg gtaccgcggtc actggatttt gggttttagga attagaaatt ttattgatag 240
 aagtatttta caaatacaaa tacatactaa gggtttctta tatgctcaac acatgagcga 300
 aacctataa gaaccctaatt tcccttatct gggaactact cacacattat tctggagaaa 360
 aatagagaga gatagatttg tagagagaga ctggtgattt ttgcgcgggg taccctaaac 420
 tgtctcacga cgttttgaac ccagattacc ctgttatccc tagtcgagcg gccgccagtg 480
 20 tgatggatat ctgcagaatt cgccctttta gatcagcaca ctggcggccg ttactagtgg 540
 atcaattcac tggccgtcgt tttacaacga ctcagagctt gacaggaggc ccgatctagt 600
 aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat ttgttttct 660
 atcgcgtatt aaatgtataa ttgcgggact ctaatcataa aaacccatct cataaataac 720
 gtcattgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata 780
 25 atcatcgcaa gaccggcaac aggattcaat cttagaagaaac ttatttgcca aatgtttgaa 840
 cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga acgcagcaag 900
 atctagagct tgggtcccgc ctacaacttc gactcccgcg ccgcgcgctg gtaccgctgg 960
 aacgcctcgt cgacgagctg cgcgacatcc tccgcgcgcg cccaactctg ctggtatccc 1020
 gcactcgaga agccatacgc atgcacaagc gtgacctcct tctccttcgc cgctcgtgcg 1080
 30 ctgcccctgc cgagcgagag gggcgacttt gtccgggtcga gaggcaggac gatccgttct 1140
 gcctcaacgc ggggtccgcc tcgtcgtgca ggtcgcaagc cgacgttggtg gcggaggacc 1200
 tcgatgcctt cgatcgttcc gtcgctcgag atggtcgggt cgaggcgcaa gcagtgcctg 1260
 aggatccgct ggaccgtctc tgggttgaca gacaagtccc agtctccac gccgtacgtc 1320
 ccgccgcaga tgacttcgcc acctggtcgg ggaatgatgt aggcgggaga agcggggtcg 1380
 35 gacgagtcca tcgtgcatcg cttgcatggg gacttgacga ggacggtttg ccgcgggatt 1440
 ggctcggcgg cttgggtcgtc gatgcccgca atcgacttgg cgccaagtcc cgtagcgttg 1500
 accaccaaat ccgcaccgtc gaacgcctgc tcaagcgacg taacggtccg tctctcaaac 1560
 gtcgcgccga gcttctgcag ctctcttgca aggtactggc agtacttttg tgcgtggacg 1620
 gagaggggtg cgtaggttac gccgatagcg ccagggtggac attcgggaaga tgggaggggg 1680
 40 cggtaatttg gcgtgatgtc cttgtaccag tgcccagaca agccgtcttc gttctgcgcg 1740
 aaccgcctcg tccccttgag ccacatggca tggcccgtcg ggaccaactc gaccacttc 1800
 ttgaaagtcg attcttccca ttttgcttgt cgaggaccgt ctgtaagcgt catgaaaggc 1860
 gtccaattcg cgccagccca tgggtgaagcg aaagtctggc tcgagacgtc ctccggcaag 1920
 tcgcgcgcga gaatatgcac gctgtagccc tcccgagcga ggatgagggc gctgctcaga 1980
 45 ccgataacgc ctgatccgag gacaacgacg cgcttctgcg agtgcattgg ccctcgacta 2040

	gagtcgagat	ccgatatcgc	ccgggctcga	gtcttttggtt	tttacttttg	ttcatgacac	2100
	tcagagactt	gagagaagca	atatatagac	ttttttttgt	tttttttttg	tggtcacgtt	2160
	tattttccta	ttggagacgg	taacgaagat	cgaacctgtg	gtggaaatga	aacaagggtg	2220
	gactagccca	cgtggtttct	tttctctgca	ttgatttggt	tttggttttt	ttgtaaagtt	2280
5	cacatcaaac	ctactaataa	ttgagaagaa	aaataaaaatc	tattgattga	ttaaaccagc	2340
	cgatgcttta	tgtctgaata	taaaaaagaa	gtgaaaaccc	cgtttaagaa	ttacaacggt	2400
	ggtttacaaa	gtatttggac	acaataaatc	caaacgaaat	aaaacaaaat	ggagaactac	2460
	caaataaaaa	acaaataaaa	aacttaaaaag	aattttattcc	attttttttc	ccgtagaatt	2520
	tattctttta	tggattcctt	aaatccatat	ttgatgcatt	ttgattcctc	ataataggta	2580
10	ataatatata	ctatgttata	gatatgtttc	taattcgtat	taacctacct	tttttttggtc	2640
	gtacgattct	acctaataat	attgaacgga	attgatgttt	tggaccactt	agaaagtatt	2700
	tttttttttg	tttgtcttag	ctgtattttca	ttaaatataa	atttaaataa	gaaatgtcat	2760
	aaataaaaatt	tgacgtatag	atttttttaa	tccattttat	gttattttaat	atttgaaatg	2820
	tgagtttggc	tcctatttta	tcttaggatg	ggttaatact	aagttttcct	taatgaatta	2880
15	tctcagagaa	actggattaa	ataaactaaa	aaatagatca	atgtgttttg	gtccggtcaa	2940
	atatcttttg	atttactatt	attggcgaaa	agaaagtctc	atatagtaaa	tcattttcct	3000
	acaagagaaa	tcaaaatttt	tgaattaaca	tggattgtat	agtttcttat	ataaccaatt	3060
	agttcgcac	aagaaaacca	aaccccaatt	aataatcaaa	cgggcttggt	aggaatattt	3120
	cattgcagct	ttcagataaa	agaaaaaac	acacactcaa	gtcttttatt	tcatttttct	3180
20	tacttgcagg	aactcaaatt	ccactttgcc	acttttcttt	acaaataaac	acaaattgtc	3240
	aatgaaacga	aatagtcttt	ttatgcaaac	actgtttgtc	ttttttcgat	cacgtttctg	3300
	attgtgacag	ccatccatat	atataggga	tgtaaaacaa	caacatgtga	agtcacatat	3360
	acgtaatgg	ttagcatagc	ttctattttc	gttgtcaata	ttagtcattc	caaaacattt	3420
	ttaagaaaaa	taaattaata	tatgtatat	cttggaacta	atgtatgtgg	aaatacagta	3480
25	acttaattat	taaacattct	aaatgcaa	atgcaaagaa	aaaaaagaaa	agaacacaac	3540
	tgaaatcaaa	gccagattca	taataattgg	ctacatgggt	gtagaatgta	gggtaacaca	3600
	acatccagaa	ttgaacactc	aaattggatg	atagatggat	aatctttaga	tacaagagaa	3660
	ttggttctct	tccattatta	acgaaaataa	agaaaaaaag	tttagcataa	aagtttgaaa	3720
	ctcaacataa	cattttgaac	ttgactcctt	cataggagtg	acatgaactg	acgaatcaca	3780
30	accgattact	tgtttgagtc	atcttccgct	ttctccacct	tcgaaatgaa	tgtgaccggt	3840
	ttcttcgggt	gctcatttac	ggcgaagtgt	aaaacatctg	gtctcgagg	acctggtagg	3900
	gataacaggg	taatctgggt	tcaaaacgtc	gtgagacagt	ttgggtgcagg	tcgaaattcg	3960
	agctcggtac	ccggtcactg	gatttttggt	ttaggaatta	gaaattttat	tgatagaagt	4020
	attttacaaa	tacaaataca	tactaagggt	ttcttatatg	ctcaacacat	gagcgaaacc	4080
35	ctataagaac	cctaattccc	ttatctggga	actactcaca	cattattctg	gagaaaaata	4140
	gagagagata	gatttgtaga	gagagactgg	tgatttttgc	gccgggtacc	gagctcggt	4200
	gcaattccc	aggctgtagc	cgacgatgg	gcgccaggag	agttgttgat	tcattgtttg	4260
	cctccctgct	gcggtttttc	accgaagtcc	atgccagtc	agcgtttttg	cagcagaaaa	4320
	gccgccgact	tcggtttg	gtcgcgagtg	aagatccctt	tcttggtacc	gccaacgcgc	4380
40	aatatgcctt	gcgaggtcgc	aaaatcggcg	aaattccata	cctgttcacc	gacgacggcg	4440
	ctgacgcgat	caaagacgcg	gtgatacata	tccagccatg	cacactgata	ctcttcactc	4500
	cacatgtcgg	tgtacattga	gtgcagccc	gctaacgtat	ccacgccgta	ttcggtgatg	4560
	ataatcggct	gatgcagttt	ctcctgccag	gccagaagtt	ctttttccag	taccttctct	4620
	gccgtttcca	aatcgccgct	ttggacatac	catccgtaat	aacggttcag	gcacagcaca	4680
45	tcaaagagat	cgctgatgg	atcgggtgtga	gcgtcgcaga	acattacatt	gacgcagggtg	4740

	atcggacgcg	tcgggtcgag	tttacgcggt	gcttccgcca	gtggcgaaat	attcccgtgc	4800
	acttgccggac	gggtatccgg	ttcgttggca	atactccaca	tcaccacgct	tgggtgggtt	4860
	ttgtcacgcg	ctatcagctc	tttaatcgcc	tgtaatgtcg	cttgctgagt	ttccccgttg	4920
	actgcctctt	cgctgtacag	ttcttttcggc	ttgttgcccc	cttcgaaacc	aatgcctaaa	4980
5	gagagggttaa	agccgacagc	agcagtttca	tcaatcacca	cgatgccatg	ttcatctgcc	5040
	cagtcgagca	tctcttcagc	gtaagggtaa	tgcgaggtag	ggtaggagtt	ggccccaatc	5100
	cagtccatta	atgcgtggtc	gtgcaccatc	agcacgttat	cgaatccttt	gccacgtaag	5160
	tccgcctctt	catgacgacc	aaagccagta	aagtagaacg	gtttgtgggt	aatcaggaac	5220
	tggtcgcctt	tactgccac	tgaccggatg	ccgacgcgaa	gcgggtagat	atcacactct	5280
10	gtctggcttt	tggctgtgac	gcacagttca	tagagataac	cttcaccogg	ttgccagagg	5340
	tgcggattca	ccacttgcaa	agtcccgcga	gtgccttgtc	cagttgcaac	cacctgttga	5400
	tccgcctcac	gcagttcaac	gctgacatca	ccattggcca	ccacctgcca	gtcaacagac	5460
	gcgtgggttac	agtcttgccg	gacatgcgtc	accacgggtga	tatcgtccac	ccaggtgttc	5520
	ggcgtgggtg	agagcattac	gctgcgatgg	attccggcat	agttaaagaa	atcatggaag	5580
15	taagactgct	ttttcttgcc	gttttcgctg	gtaatcacca	ttcccggcgg	gatagtctgc	5640
	cagttcagtt	cgttgttcac	acaaacgggt	atacctgcac	atcaacaaat	tttggtcata	5700
	tattagaaaa	gttataaatt	aaaatataca	cacttataaa	ctacagaaaa	gcaattgcta	5760
	tatactacat	tctttttattt	tgaaaaaaat	atttgaaata	ttatattact	actaattaat	5820
	gataattatt	atatatatat	caaaggtaga	agcagaaact	tacgtacact	tttcccggca	5880
20	ataacatacg	gcgtgacatc	ggcttcaaatt	ggcgtatagc	cgccctgatg	ctccatcact	5940
	tcctgattat	tgaccacac	tttgccgtaa	tgagtgaccg	catcgaaacg	cagcacgata	6000
	cgctggcctg	cccaaccttt	cggtataaag	acttcgcgct	gataccagac	gttgcccgcg	6060
	taattacgaa	tatctgcac	ggcgaactga	tcgttaaaac	tgccctggc	agcaattgcc	6120
	cggtcttctt	gtaacgcgct	ttcccaccaa	cgctgaccaa	ttccacagtt	ttcgcgatcc	6180
25	agactgaatg	cccacagggc	gtcgagtttt	ttgattttcac	gggttggggg	ttctacagga	6240
	cgtaccatgg	tcgatcgact	ctagactagt	ggatccgata	tcgcccgggc	tcgactctag	6300
	atgaaatcga	aattcagagt	tttgatagtg	agagcaaaga	gggacggact	tatgaggatt	6360
	tcgagtattt	caagagatgg	tacttggtga	tcggacggct	acgatgatct	cgatttggtt	6420
	aatccagtat	ctcgcgggtg	atggagttat	ggtaggggtta	atgggtcaatt	tcactctaag	6480
30	gtagagaatg	atgtaattag	ataagaatct	tgagatactg	gttttagattg	gatgagtgtg	6540
	gggtccatct	tatcttgata	agtggatggg	tttttagagac	acagtgaata	ttagccaatc	6600
	gaagttccat	atcaccatca	tcactctgtat	aattttgttt	ttttggaaga	taataatgat	6660
	tgaatttttg	gtagattttta	ttttttcatta	tttaccttgt	atgttgagtg	gtcttcaaatt	6720
	tattgaacgt	gacagattca	caagaaagta	gatttttttat	aaatgaaatt	ttactttattt	6780
35	taaagggtatc	tctattttaat	ttctttttgtt	tatgggttgct	tgctcagcatt	tgacttgacg	6840
	tttcatgctc	atagtcatat	acgttattct	aggctttttt	gaatatctta	ttactttttt	6900
	cgtaatacaa	ttttataatt	ttatcaaagt	tatacaacta	taactaaaat	taggggttttc	6960
	tacaaaacaa	aaaaatcttc	taattttttt	tgttgtagcc	agtttactcg	taagttacaa	7020
	aaaaatacaa	atgaaccac	atgtattatg	cgtttaacta	ggattaccat	gtactttcat	7080
40	gtactcaatt	caccctatac	tctttttttt	tttttttcta	gttccacca	atctataaaa	7140
	ttctgtccat	ttgaccaaatt	tcaatttaatt	tctgtaattg	cgatttaaaa	ttaatattac	7200
	atgttcaacta	tttctcgatt	tgagggaacc	cgagtttaaa	tatgataaaa	atgttgaccc	7260
	atcactacaa	atatgttata	gtttataact	aatagtgggtg	tttttgggga	taattgatga	7320
	attaagtaaa	catgattctt	cttatgaagt	tgattgagtg	attattgtat	gtaaacctat	7380
45	gtgattgatg	ttattgggtg	attgagtgat	tattgtatta	gtatgtaagc	aaagatgatt	7440

	gttcttatga	ggtaatttgt	tactcattca	tccttttgca	tatgagaaat	tgtgttagcg	7500
	tacgcaaaac	aatagagAAC	ataaaagata	tgtgtattta	tttaagggtga	cttttgttaa	7560
	tgatattgta	gtatctatac	atttatatat	aacttggtga	atttgagtat	aagctatcag	7620
	gatccggggg	atcctctaga	gtcagaggtac	cgagctcgaa	ttcactggcc	gtcgtttttac	7680
5	aacgactcag	tactgcttgg	taataattgt	cattagattg	tttttatgca	tagatgcact	7740
	cgaaatcagc	caattttaga	caagtatcaa	acggatgtta	attcagtaca	ttaaagacgt	7800
	ccgcaatgtg	ttattaagtt	gtctaagcgt	caatttggtt	acaccacaat	atatacctgcc	7860
	accagccagc	caacagctcc	ccgaccggca	gctcggcaca	aaatcaccac	gcgtctaaaa	7920
	aggtgatgtg	tatttgagta	aaacagcttg	cgtcatgcgg	tcgctgcgta	tatgatgcga	7980
10	tgagtaaata	aacaaatac	caaggggaac	gcatgaagg	tatcgctgta	cttaaccaga	8040
	aaggcgggtc	aggcaagac	accatcgcaa	cccatctagc	ccgcgccctg	caactcgccg	8100
	gggccgatgt	tctgttagtc	gattccgata	cccaggggcag	tgcccgcgat	tgggcggccg	8160
	tgccgggaaga	tcaaccgcta	accgttgctg	gcatcgaccg	cccgacgatt	gaccgcgacg	8220
	tgaaggccat	cggccggcgc	gacttcgtag	tgatcgacgg	agcgccccag	goggcggact	8280
15	tggctgtgtc	cgcgatcaag	gcagccgact	tcgtgctgat	tccggtgcag	ccaagccctt	8340
	acgacatatg	ggccaccgcc	gacctgggtg	agctgggtta	gcagcgcatt	gaggtcacgg	8400
	atggaaggct	acaagcggcc	tttgctcgtg	cgcggggcgat	caaaggcacg	cgcatacggc	8460
	gtgaggttgc	cgaggcgctg	gccgggtacg	agctgcccac	tcttgagtcc	cgtatcacgc	8520
	agcgcgtgag	ctaccagggc	actgcgcgcc	ccggcacaac	cgttcttgaa	tcagaaccgc	8580
20	agggcgacgc	tgcccgcgag	gtccagggcgc	tggccgctga	aattaaatca	aaactcattt	8640
	gagttaatga	ggtaaagaga	aaatgagcaa	aagcacaac	acgctaagt	cggccgctcc	8700
	gagcgcacgc	agcagcaagg	ctgcaacggt	ggccagcctg	gcagacacgc	cagccatgaa	8760
	gcgggtcaac	tttcagttgc	cggcggagga	tcacaccaag	ctgaagatgt	acgcggtacg	8820
	ccaaggcaag	accattaccg	agctgctatc	tgaatacatc	gcgcagctac	cagagtaaata	8880
25	gagcaaatga	ataaatgagt	agatgaattt	tagcggctaa	aggaggcggc	atggaaaatc	8940
	aagaacaacc	aggcaccgac	gccgtggaat	gccccatgtg	tggagggaac	ggcgggttggc	9000
	caggcgtaag	cggctgggtt	gtctgccggc	cctgcaatgg	cactggaacc	ccaagcccgc	9060
	aggaatcggc	gtgagcggtc	gcaaaccatc	cggcccggta	caaatacggc	cggcgcctggg	9120
	tgatgacctg	gtggagaagt	tgaaggccgc	gcaggccgcc	cagcggcaac	gcatcgaggc	9180
30	agaagacgcc	ccggtgaatc	gtggcaaggg	gccgctgata	gaatccgcaa	agaatcccgc	9240
	caaccgccgc	cagccgggtg	gccgtcgatt	aggaagccgc	ccaagggcga	cgagcaacca	9300
	gatttttttc	ttccgatgct	ctatgacgtg	ggcaccgcgc	atagtcgcag	catcatggac	9360
	gtggccggtt	tccgtctgtc	gaagcgtgac	cgacgagctg	gcgaggtgat	ccgctacgag	9420
	cttccagacg	ggcacgtaga	ggtttccgca	gggccggccg	gcatggccag	tgtgtgggat	9480
35	tacgacctgg	tactgatggc	ggtttcccat	ctaaccgaat	ccatgaaccg	ataccgggaa	9540
	gggaagggag	acaagcccgc	ccgcgtgttc	cgtccacacg	ttgcggacgt	actcaagttc	9600
	tgccggcgag	ccgatggcgc	aaagcagaaa	gacgacctgg	tagaaacctg	cattcgggtta	9660
	aacaccacgc	acgttgccat	gcagcgtacg	aagaaggcca	agaacggccg	cctggtgacg	9720
	gtatccgagg	gtgaagcctt	gattagccgc	tacaagatcg	taaagagcga	aaccgggcgc	9780
40	ccggagtaca	tcgagatcga	gctagctgat	tggatgtacc	gcgagatcac	agaaggcaag	9840
	aaccgcgacg	tgctgacggt	tcaccccgat	tactttttga	tcgatcccgc	catcggccgt	9900
	tttctctacc	gcctggcacg	ccgcgccgca	ggcaaggcag	aagccagatg	gttgttcaag	9960
	acgatctacg	aacgcagtgg	cagcgcgcga	gagttcaaga	agttctgttt	caccgtgcgc	10020
	aagctgatcg	ggtcaaatga	cctgccggag	tacgatttga	aggaggaggc	ggggcaggct	10080
45	ggcccgatcc	tagtcatgcg	ctaccgcaac	ctgatcgagg	gcgaagcatc	cgcgggttcc	10140

	taatgtacgg	agcagatgct	agggcaaatt	gccctagcag	gggaaaaagg	tcgaaaaggt	10200
	ctctttcctg	tggatagcac	gtacattggg	aacccaaagc	cgtacattgg	gaaccggaac	10260
	ccgtacattg	ggaacccaaa	gccgtacatt	gggaaccggt	cacacatgta	agtgactgat	10320
	ataaaaagaga	aaaaaggcga	tttttccgcc	taaaactctt	taaaacttat	taaaactctt	10380
5	aaaacccgcc	tggcctgtgc	ataactgtct	ggccagcgca	cagccgaaga	gctgcaaaaa	10440
	gcgctacccc	ttcggctcgct	gcgctcccta	cgcgccgccc	cttcgcgctc	gcctatcgcg	10500
	gcctatcgcg	tgtgaaatac	cgcacagatg	cgtgaaggaga	aaataaccgca	tcaggcgctc	10560
	ttccgcttcc	tcgctcactg	actcgctgcg	ctcggctcgtt	cggctgcggc	gagcggtatc	10620
	agctcactca	aaggcggtaa	tacggttatc	cacagaatca	ggggataacg	caggaaagaa	10680
10	catgtgagca	aaaggccagc	aaaaggccag	gaaccgtaaa	aaggccgcgt	tgctggcggtt	10740
	tttccatagg	ctccgcccc	ctgacgagca	tcacaaaaat	cgacgctcaa	gtcagagggtg	10800
	gcgaaacccg	acaggactat	aaagatacca	ggcgtttccc	cctggaagct	ccctcgctcg	10860
	ctctcctggt	cgcacctgc	cgcttaccgg	atacctgtcc	gcctttctcc	cttcgggaag	10920
	cgtggcgctt	tctcatagct	cacgctgtag	gtatctcagt	tcgggtgtagg	tcgttcgctc	10980
15	caagctgggc	tgtgtgcacg	aacccccgt	tcagccccgac	cgtgcgcct	tatccggtaa	11040
	ctatcgtctt	gagtccaacc	cggtaagaca	cgacttatcg	ccactggcag	cagccactgg	11100
	taacaggatt	agcagagcga	ggtatgtagg	cgggtgctaca	gagttcttga	agtgggtggc	11160
	taactacggc	tacactagaa	ggacagtatt	tggtatctgc	gctctgctga	agccagttac	11220
	cttcggaaaa	agagttggta	gctcttgatc	cggcaaaaaa	accaccgctg	gtagcgggtg	11280
20	tttttttgtt	tgcaagcagc	agattacgcg	cagaaaaaaa	ggatctcaag	aagatccttt	11340
	gatcttttct	acgggggtctg	acgctcagtg	gaacgaaaac	tcacgttaag	ggatttttgt	11400
	catgcatgat	atatctccca	atttgtgtag	ggcttattat	gcacgcttaa	aaataataaa	11460
	agcagacttg	acctgatagt	ttggctgtga	gcaattatgt	gcttagtgca	tctaacgctt	11520
	gagttaagcc	gcgcccgcga	gcggcgctcg	cttgaacgaa	tttctagcta	gacattatct	11580
25	gccgactacc	ttgggtgatct	cgcctttcac	gtagtggaca	aattcttcca	actgatctgc	11640
	gcgcgaggcc	aagcgatctt	cttcttgctc	aagataagcc	tgtctagctt	caagtatgac	11700
	gggctgatac	tgggcccggc	ggcgctccat	tgcccagtcg	gcagcgacat	ccttcggcgc	11760
	gattttgccc	gttactgcgc	tgtaccaaatt	gcgggacaac	gtaagcacta	catttcgctc	11820
	atcgccagcc	cagtcggggc	gcgagttcca	tagcgttaag	gtttcattta	gcgcctcaaa	11880
30	tagatcctgt	tcaggaaccg	gatcaaagag	ttcctccgcc	gctggaccta	ccaaggcaac	11940
	gctatgttct	cttgcttttg	tcagcaagat	agccagatca	atgtcgatcg	tggtgggctc	12000
	gaagatacct	gcaagaatgt	cattgcgctg	ccattctcca	aattgcagtt	cgcgcttagc	12060
	tggataacgc	cacggaatga	tgtcgtcgtg	cacaacaatg	gtgacttcta	cagcgcgagg	12120
	aatctcgctc	tctccagggg	aagccgaagt	ttccaaaagg	tcgttgatca	aagctcgccg	12180
35	cgttggtttca	tcaagcctta	cggtcaccgt	aaccagcaaa	tcaatatcac	tgtgtggctt	12240
	caggccgcca	tccactgcgg	agccgtacaa	atgtacggcc	agcaacgtcg	gttcgagatg	12300
	gcgctcgatg	acgccaacta	cctctgatag	ttgagtcgat	acttcggcga	tcaccgcttc	12360
	ccccatgatg	tttaactttg	ttttagggcg	actgccctgc	tgcgtaacat	cgttgctgct	12420
	ccataacatc	aaacatcgac	ccacggcgta	acgcgcttgc	tgcttggatg	cccgaggcat	12480
40	agactgtacc	ccaaaaaaac	agtcataaca	agccatgaaa	accgccactg	cgttccatg	12539